Undocumented Orphan Wells Program Annual Report

February 2024

Govert, Andrew; Viswanathan, Hari; McNamara, Elizabeth

List of Acronyms

AGL	Above ground level
AVIRIS	Airborne Visible/Infrared Imaging Spectrometer
BIL	Bipartisan Infrastructure Law
CATALOG	Consortium Advancing Technology for Assessment of Lost Oil & Gas Wells
CALGEM	California Geologic Energy Management Division
CCUS	Carbon Capture Utilization and StorageCH_4 Methane
DC	Direct Current
DCE	Direct Current Electrification
DEP	Department of Environmental Protection
DOE	U.S. Department of Energy
DOI	U.S. Department of the Interior
DOW	Documented Orphaned Wells
EDX	Energy Data eXchange
EM	Electromagnetic
EMD	Empirical Mode Decomposition
EM-TDR	Electromagnetic Time Domain Reflectometry
EPA	U.S. Environmental Protection Agency
FECM	Fossil Energy and Carbon Management
FOA	Funding Opportunity Announcement
GHG	Greenhouse gas
GIS	Geographical Information System
GPR	Ground Penetrating Radar
GWR	Geographically Weighted Regression
HiFEM	Hierarchical Finite Elements
HIFLD	Homeland Infrastructure Foundation Level Data
HTMC	Historical Topographic Maps Collection
IMF	Intrinsic Mode Functions
IOGCC	Interstate Oil and Gas Compact Commission
IOW	Inadequately Documented Orphan Wells
IP	Induced Polarization
LANL	Los Alamos National Labs
LBNL	Lawrence Berkley National Laboratory
LiDAR	Light Detection and Ranging
LLM	Large Language Model
LLNL	Lawrence Livermore National Labs
MERP	Methane Emissions Reduction Program
ML	Machine Learning
NAIP	National Agriculture Imagery Program
NETL	National Energy Technology Laboratory
NORM	Naturally Occurring Radioactive Material
O&G	Oil and gas
OCR	Optical Character Recognition
OLS	Ordinary Least Squares
OWPO	Orphaned Well Program Office
OW	Orphan Wells
PI	Principal Investigator
R&D	Research and Development
RF	Random Forest
SNL	Scandia National Labs
SNR	Signal to Noise Ratio
TDF	Time Domain Reflectometry
TEM	Time- Domain Electromagnetic
TPR	True Positive Rate
UAV	Unmanned Aerial Vehicle
UOW	Undocumented Orphaned Wells
UOWP	Undocumented Orphaned Wells Program
USDA	U.S. Department of Agriculture
USGS	U.S. Geological Survey
WP	Work Packages
WT	Wavelet Transform

Table of Contents

1. Mission and Objectives **8**
Executive Summary 8
Program Summary 8
2. Summary of Key Work Packages **10**
2.1 Work Package 1: Methane Detection and Quantification 10
Work Package 2: Well Identification 11
Work Package 3: Sensor Fusion and Data Integration with Machine Learning 11
Work Package 4: Characterization 12
Work Package 5: Integration and Best Practices 13
Work Package 6: Data Management 13
3. Work Package 1- Methane Detection and Quantification **14**
3.1 Section Summary 14
3.2 Activity 15
3.3 Motivation and Background 21
3.4 Data Summary 22
3.5 Future Work 22
4. Work Package 2- Well Identification **23**
4.1 Section Summary 23
4.2 Activity 25
4.3 Motivation and Background 31
4.4 Data Summary 34
4.5 Future Work 34
5. Work Package 3- Sensor Fusion and Data Integration with Machine Learning **35**
5.1 Section Summary 36
5.2 Activity 36
Key Accomplishments for each activity are summarized below: 36
5.3 Motivation and Background 37
5.4 Data Summary 37
5.5 Future Work 58
6 Work Package 4- Characterization **58**
6.1 Section Summary 59
6.2 Activity 60
6.3 Motivation and Background 73

6.4 Data Summary 74
6.5 Future Work 74
7 Work Package 5- Integration and Best Practices 75
7.1 Section Summary 75
7.2 Activity 75
7.3 Motivation and Background 76
7.4 Data Summary and Accomplishments 78
7.5 Future work 86
8 Work Package 6- Data Management 86
8.1 Section Summary 87
8.2 Activity 87
8.3 Motivation and Background 88
8.4 Data Summary 89
8.5 Future Work 89
9 Work Package 7- Data Extraction 89
9.1 Section Summary 89
9.2 Activity 90
9.3 Motivation and Background 90
9.4 Data Summary 91
9.5 Future Work 91
10 Work Package 8- Composite Database and Dashboard 92
10.1 Section Summary 92
10.2 Activity 92
10.3 Motivation and Background 92
10.4 Data Summary 93
10.5 Future Work 93
11 Work Package 9- Field Teams 94
11.1 Section Summary 94
11.2 Activity 94
11.3 Motivation and Background 95
11.4 Data Summary 96
11.5 Future Work 96
12 Appendix A - References 96

LIST OF EXHIBITS

Figure 1.1: BIL funded programs targeting methane mitigation and reduction from historical oil & gas wells. Due to the nature of the challenge these bills address there are many points of overlap. Federal agencies are working to collaborate across these programs to provide for the most efficient use of funds. 9
Figure 1.2: Demonstrating the flow of information from CATALOG to the DOI OWPO's $4.7 billion orphan well plugging program. 10
Figure 3.1: Example of equipment assessment –methane concentration mapping and emission source identification on a Digital Twin of a compressor inside an operating compressor station. Left: leak identification, Right: post-repair confirmation of emissions mitigation. 16
Figure 3.2: Experimental setup for our plume monitoring and gaussian plume model analysis to infer emissions based on wind statistics and methane concentrations measured downwind of the point source. 17
....18
Figure 3.3: (Left panel) Measurement of downwind methane concentration (1 m) from the point source, aligned with the center line of the plume as a function of the leak rate. The red line is a fit through the measured methane concentrations. Methane concentration variability is caused by the change in the position of the plume as shown in the Gaussian plume profile (right panel). 18
Figure 3.4: Gas concentrations and composition, and flow rates measured at orphan wells in Hillman State Park, PA. Upper left: Xplorobot sensor; Upper right: RMLD sensor; Lower left: staff from PA-DEP collecting a sample for gas composition analysis; Lower right: NETL and LBNL Team. 19
Figure 3.5: Field set up, wind data and plume model used in our Hobbs, NM orphan well campaign in April 2023. 20
Figure 3.6: Top panel: measured emissions flow rate of the venting well measured by WellDone using a Ventbuster flow meter. Middle panel: Time series of the wind direction as measured by the Trisonica 3-D anemometer at 7.5 m downwind of the well site. Bottom panel: Time series analysis of the methane enhancement measured by the four methane spectroscopic instruments downwind of the venting Foster 1.1S well at 7.5 m, 15 m, 22.5 m, and 47 m. Annotations mark when Foster 1.1S vent was opened and when the vent was closed. 20
Figure 3.7: Example of short-term variability in methane emissions from several wells in Hillman State Park, PA. 21
Table 4.1: Selection criteria for ground, UAV and helicopter surveys for magnetic and electromagnetic data collection. 24
Table 4.2: Selection criteria for magnetic and electromagnetic data collection 24
Figure 4.1: Drone based EM system and example dataset for steel casing identification. (A) Drone based EM system with autonomous flight planning; (B) flight path for test flight with red dot showing the location of the well; (C) EM anomaly identified from the EM data that is associated with the well; (D) A linear view of the EM signal associated with the wellbore in comparison with the background. 27
Figure 4.2. Drone TEM responses at 40 m flight heights for non—metallic, clayey (left) and metallic (right) wells. Homogeneous half space response (HHS) is shown for reference. 28
Figure 4.3: Surrogate 2-D numerical model to determine sensitivity of drone-GPR to well identification; (a) Dielectric map of the 2-D model, (b-d) electric field snapshots at 12 ns, 24 ns and 36 ns showing waves hitting the ground and scattering from the well and reflecting back to the antenna, (e) radargram can detect the presence of the well along with source contribution and ground reflection, (f) sensitivity of GPR signal to soil conductivity (GPR suitability index ranging from 1-5), (g) sensitivity of GPR signal to lack of casing. 29

Figure 4.4: GPR system (250 MHz-1 GHz) integrated on a cart platform at LLNL scanning over a concrete platform (Frequency chosen strategically for deep penetration (1 meter) with < 1 cm resolution); (https://www.sensoft.ca/products/pulseekkopro/overview-pulseekko/) 30
Figure 4.5. HiFEM validation exercise demonstrating the equivalence of volume-based (expensive) and edge-based (cheap) well representations for a top-casing stimulation scenario. 31
Figure 4.6. Modular design architecture of HiFEM multi-physics ecosystem whereby meshing and data visualization workflows communicate with HiFEM through standardized I/O progams. 31
....32
Figure 4.7: Heat map of well density (red-high, green-low, white-none) filtered by GPR suitability index (1-most suitable, 5-least suitable). Most of the well locations lie in regions with moderate to high GPR suitability (middle panel). 32
Figure 4.8. In the hierarchical materials properties framework (Weiss, 2017), electrical properties can be attributed to volumes (left), facets (middle) and edges (right) of the underlying discretization, thus minimizing the cost of representing thin features within a given simulation. 33
Figure 4.9. Numerical simulation of the DC response of a section within the Kern River oilfield [*] whereby top-casing electrification of a single well head (blue region, left) results in secondary electrification of nearby wells and top-side infrastructure such as the distribution pipelines and storage tanks shown here. 34
Table 5.1: WP3 Activity Leads (Blue) and Activity Partners (Tan) 37
Figure 5.1: Maps showing wells organized by spud date. (a) About ⅔ of the wells in the CALGEM All wells database do not have information on spud dates. (b) Unsupervised spatial clustering of nearby wells spudded within the same decade indicate potential patterns in well drilling. 38
Figure 5.2: Well count difference per 2-square mile between the IHS aggregated dataset and CalGEM..39
Table 5.2 Value Proposition Matrix for Sensors 40
Figure 5.3 Procurement and Initial Validation for Sensors Selected..41
Figure 5.4: Image constructed by Lidar scans, coloring done by elevation, green lower in elevation and red higher in elevation. 41
Figure 5.5: A range of environmental geospatial data are considered in creating a GIS model describing the likelihood of O&G wells being present. 42
Figure 5.6: Heat map of O&G wells on record on Oklahoma (purple, low density; yellow, high density). White polygon delineates Lincoln County. 43
Figure 5.7: O&G wells on Lincoln County are highly spatially clustered (statistical distribution curve). No clear temporal clustering is observed. Peach dots represent 1 well in a 30m x 30m area; pink dots represent 2 or more wells in a 30m x 30m area. (Well Data Source: Risk Based Data Management System for the Oklahoma Corporation Commission.) 43
Figure 5.8: OLS estimated well density: Yellow, overestimate, white, good estimate, brown, underestimate. 44
Figure 5.9: Bar chart of variable importance in Random Forest Regression. Importance (values relative to this specific model) is a measure of how much outputs change when a variable is omitted from the model. Magnetics is the aeromagnetic survey data. 45
Figure 5.10: Example of census records with O&G industry-related occupations....................................46
Figure 5.11: Examples of NAIP images that correspond with O&G wells from Boutot et al., (2022) database. 47
Figure 5.12: Image classification is highly dependent on the amount of tree canopy in imagery. Dense tree cover does not have a clear well signature. 47
Figure 5.13: Examples of historic documents and data sources that can portray or be used to confirm undocumented well locations, including A) USGS quadrangle maps, B) aerial photography, C) LIDAR surveys, and D) farm line maps. 48

Figure 5.14: A) Density of wells with spud dates greater than 50 years ago as provided by the Co2-Lcoate, version 1 database. B) Breakdown of wellbore records based off status information provided, and if provided how those might breakdown and related to orphaned, marginal, or other types of wells. C) Data acquisition progress for national well database to spotlight states where public wellbore data has been collected and completed, is in progress, or has been identify for future needs. 50
Figure 5.15: a) Inset of historical topographic maps in Lincoln County, OK with detected wells highlighted by blue circles, documented wells in purple and UOWs in red. Blue numbers are unique ids, while burgundy values represent the distance in meters of a UOW from the closest documented well. b) The same area in a) shown from Google Maps. UOWs and documented wells are highlighted by the same symbols as in (a). In (c) a zoomed version of (b) allows one to visually distinguish the pump jacks. 51
Figure 5.16: a) Formulas to calculate the True Positive Rate and Precision. b) Representative image of a circular map well compared to a template circle. c) Inset of HTMC map with detected wells highlighted by blue square and the false positives highlighted by red squares. d) ArcGIS model to create labelled training sets by creating a buffer around the map symbol and extracting the features. e) Labelled circular map symbols used to train and test the model. f) Template matching model metrics. 53
Figure 5.17: Estimated number of UOWs per county, including ER wells and potential underreporting in the past (a) in Pennsylvania during the entire state history 1859-2023 , and (b) in Oklahoma during the entire state history 1891-2023. 54
Figure 5.18: (Top) Diagram of the Process: Different types of signals (like methane and magnetic) are combined and fed into a transformer model. This model is trained to understand how these signals can show where wells might be in a certain area. 55
(Bottom) Results: When we use both types of signals, the model predicts all the wells in our dataset with a high degree of accuracy. 55
Figure 5.19: Images of methane plumes are used to estimate leak rates. The figure title compares the true leak rates with the predicted leak rates. 56
Figure 5.20: Synthetic magnetic anomaly of a steel casing. The white circle is the location of a vertical steel casing. Color indicates the intensity of the magnetic anomaly. 57
Figure 5.21: Parameters affecting magnetic anomaly of a steel casing. (a) magnetometer altitude. (b) Casing length. (c) Steel casing length. (d) Magnetic susceptibility of the steel casing. 58
Figure 6.1: The operating principle of Electromagnetic Time Domain Reflectometry (EM-TDR) under development for WP4. 61
Figure 6.2: The conceptual setup of the EM-TDR system showing the major electrical components involved and the signal transmission pathway. 62
Figure 6.3: Example results of the EM-TDR tests on tubing with engineered damage. A-B: Engineered damaged made on metallic tubing acquired from an operator with a circular thinning on the tubing at 0.5-inch diameter and 25% wall thickness loss; C: EM-TDR trace for the 2GHz signal indicating the key features corresponding to those on the tubing; D: Wavelet transformed EM-TDR signals across the frequency bandwidth tested from 100M Hz to 2 GHz. 63
Figure 6.4: Numerical modeling and simulation of the EM fields for O&G wells due to corrosion damages. (a) The model wellbore diagram used in the simulation where two cross sections (b and c) are used for the 2D simulation of the EM responses due to corrasion. The top row of colored b figures (b1-3, c1-3) represents the 3-casing layer borehole cross section showing the corrosion damage feature in the red box (b1), the potential field distribution when the casing is energized (b2), and the changes in the potential field due to the corrosion damage (b3). The bottom c row shows the same features for the damage at a deeper location (c) shown on the borehole diagram where only 2 layers of casing are present. 64
Figure 6.5: O&G wells in Oklahoma used for the initial field testing of the prototype EM-TDR system. Wells # 1-3 are located in Holdenville, OK and Wells # 4-5 are located on the University of Tulsa campus.

Wells #1-4 are 5 ½ inch diameter wellbore wells while Well #5 is a 12 wellbore well. All wells are at the depths ~ 3000 ft. 65
Figure 6.6: Example EM-TDR dataset from an O&G wellbore in Holdenville, OK. A: The test borehole; B: EM-TDR traces for frequencies at 770 MHz, 500 MHz, 300 MHz and 150 MHz. The identified features are marked as dashed red lines. 65
Figure 6.7: 2d mesh used to mimic airborne GPR acquisition. The mesh contains 83,812 unstructured quadrilateral spectral- elements. GPR source is a Ricker function of central frequency 500 MHz. A schematic drone is displayed showing altitude of 1 m and direction of the acquisition. 67
Figure 6.8: Zoom in showing the location of the 11 source- antenna pair locations with respect to the well. Sources are displayed as yellow crosses and receiver antenna as green squares. 67
Figure 6.9: Zoom in snapshots of the electric field at 17.5 ns for source-antenna pair at location 4, offset from the well location by 2 m. Panel (a) depicts the wavefield with the presence of a well, while panel (b) depicts the wavefield without a well. 67
Figure 6.10: Stacked radargrams with and without well, and their difference. Raw radargrams present two strong signals: direct and surface reflected. The difference highlights the presence of the well in the classical form of a parabolic signal. 68
Figure 6.11: Diagram of the well we used to test acoustic methods. 69
Figure 6.12: Spectra for auto-correlation with horizontal axis converted to velocity if the travel distance is the two-way distance through the well casing. Red, green, blue indicate east, north, vertical, respectively. 70
Figure 6.13: Modeled experimental setup for "top-casing" method of wellbore integrity assessment (Wilt, 2016; MacLennan et al., 2018). 71
Figure 6.14: Distribution of electrostatic potential in the subsurface (left) arising from top-casing excitation for a fully in-tact well (left, top) with uniform electrical conductivity, and a well with a mid-depth region of reduced conductivity due to "damage" from corrosion, breakage, etc. Surface profiles of radial electric field (right) show that the damaged well is distinguishable from the intact well through differences in its near-offset curvature and amplitude. 71
Figure 6.15: Profiles of radial electric field (total field, left; relative field, right) on Earth's surface over the horizontal section of a deviated well show the effect of the damage location on measured response *(Beskardes et al., 2021).* 72
Figure 6.16: Inversion results for damage zone quantification using surface-based E-field measurements and top-casing excitation for mid-well (top row) and bottom-well (bottom row) damage scenarios. Constraining the free parameters in the inversion (4th column) to the regions where the damage is believed to lie allows for acceptable recovery of the damage zone conductivity (3rd, 5th and 6th columns). 72
Figure 6.17: Inversion results for constraining both well damage magnitude and location using a combination of data collected on both the surface and inside a neighboring "observation well". Damage location is reasonably constrained, regardless of discretization coarseness (top and bottom rows) and robust with respect to bounding estimates of observational noise (middle row). 73
Figure 7.1: From Dubey et al. (2023). Methane emission flow rates are shown as a function of measured concentration 2m from the source and wind speed. Horizontal red lines are supplied for context; 1 g/hr is the methane reporting requirement for the BIL 2021, EPA 17 g/hr is the limit of detection using optical gas imaging to detect methane leaks. 79
Figure 7.2: From Reeder et al. (2023). Data compilation at a project site in New York State. 80
Figure 7.3: Excerpt of results from an aerial magnetic survey at Teapot Dome Oilfield, WY (from Hammack et al., 2023). Each map shows magnetic intensity at different aircraft heights and inline grid spacing during the survey. Magnetic anomalies that indicate well locations lose resolution as height and grid spacing increase. 81

Figure 7.4: Documented orphaned wells as of IIJA enactment, from IOGCC (2021). 82
Figure 7.5: Field testing and demonstration focus areas for the five national labs participating in the CATALOG program. 83
Table 7.1: Summary of Key Terms associated with CATALOG Project. Key synonyms from Boutot et al (2022) are included, especially for states being considered for project field testing activities.................84
Table 7.2: Milestones for WP5...85
....88
Figure 8.1: Summary statistics about users and viewers of the CATALOG website since it's public release in May 2023. 88
Figure 9.1: Information about inadequately documented oil & gas wells can be extracted from historical records using large language models. Here, we show information about the location and depth of a well extracted from a scanned document with complicating factors like stamps and handwritten notes. These forms sometimes have errors (such as the longitude having the wrong sign – indicating the wrong hemisphere of the globe) that large language models can automatically correct. 90
Table 9.1: Activity Leads (Blue) and Activity Partners (Tan)..90
Figure 10.1: Comparison of wellbore records obtained for version 1 of the CO2-Locate database (available for download on EDX, https://edx.netl.doe.gov/dataset/co2-locate) and additional wellbore resources added for version 2 of the database, with a public release anticipated in Spring 2024. 93
Figure 11.1: Areas of interest. WP9 is planning field campaigns at the Four Corners (Chula Oil Fields), NM, Osage Nation, OK, and Stonewall Jackson State Park, WV. Central Valley, CA is another potential area for WP9. Remaining stars have areas previously surveyed by NETL (pre-UOWP-CATALOG). 95
Figure 11.2: Osage County, OK. Well density (BIA O&G well database) and two priority properties (black polygons) identified by Osage Nation and the Osage Mineral Council. Focus areas (green polygons) on Osage Nation Ranch (west) and Lost Creek Ranch (east) are planned for the first of multiple field deployments. The orange dots are the results of WP3's Historic Topographic Map Classification algorithm for Osage County...95

1. Mission and Objectives

Executive Summary

The Undocumented Orphan Wells Program (UOWP) was formed through language in the Bipartisan Infrastructure Law (BIL) signed by President Biden on November 15, 2021. The BIL directs the U.S. Department of Energy (DOE) to collaborate with the Interstate Oil and Gas Compact Commission (IOGCC) and the U.S Department of Interior (DOI) Orphaned Well Program Office (OWPO) to develop a program focused on reducing the impact of undocumented orphaned wells (UOWs). The BIL allocated $30 million dollars to the effort. The focus of the UOWP is to establish a consortium of National Labs dedicated to developing techniques, technologies and best practices for the identification and characterization of undocumented orphan wells, and to share the results of this work with federal, state, and tribal partners.

In September of 2022 the Consortium Advancing Technology for Assessment of Lost Oil & Gas Wells (CATALOG) was initiated and consisted of six work packages (WP): WP 1 Methane Detection and Quantification, WP 2 Well Identification, WP 3 Sensor Fusion and Data Integration with Machine Learning, WP 4 Characterization, WP 5 Integration and Best Practices, and WP 6 Data Management. Five National Labs are collaborators: Los Alamos National Labs (LANL), Scandia National Labs (SNL), Lawrence Livermore National Labs (LLNL), Lawrence Berkley National Labs (LBNL) and the National Energy Technology Laboratory (NETL).

The Fiscal Year 2023 Federal budget included language to provide annual funding support to the UOWP program. In FY23 the program was directed to spend up to $10 million in annual appropriations on the effort. Of this, $5 million was used to expand CATALOG with three new work packages: WP 7 Records Data Extraction, WP 8 Wells Database, and WP 9 Field Teams. These WPs are directed outgrowths of the original WPs, and cover areas where additional funds can significantly boost the success of the overall program and are easily distinguishable from BIL funded work. The funding for these WPs was released in August of 2023 and work is underway. The remaining FY23 funding is allotted for an upcoming Funding Opportunity Announcement (FOA), as requested in the FY23 budget language.

This report serves as the primary summary of all UOWP activity over the year. It has been designed to meet the reporting requests provided in the BIL language and will set the standard for the yearly reports throughout the life of the program. While FY2023 funded activities are not required to be reported in this report, their inclusion provides a much more complete picture of the UOWP.

Program Summary

On November 15, 2021, the Bipartisan Infrastructure Law (BIL) was signed by President Biden. Within the BIL was language requesting the U.S. Department of Energy (DOE) collaborate with the Interstate Oil and Gas Compact Commission (IOGCC) and the Department of Interior (DOI) Orphaned Well Program Office (OWPO) to develop a program focused around reducing the impact of undocumented orphaned wells (UOWs). To meet this request the DOE formed the Undocumented Orphan Wells Program (UOWP). This program is one of three “end of life” methane mitigation programs targeting legacy wells across the nation (Figure 1.1). Per the BIL language, the UOWP is to advance and share technologies, techniques, and best practices for identifying and characterizing UOWs. To achieve these goals UOWP set up the Consortium

Advancing Technology for Assessment of Lost Oil & Gas Wells (CATALOG) to provide support to state, federal and tribal entities in addressing UOWs across the nation.

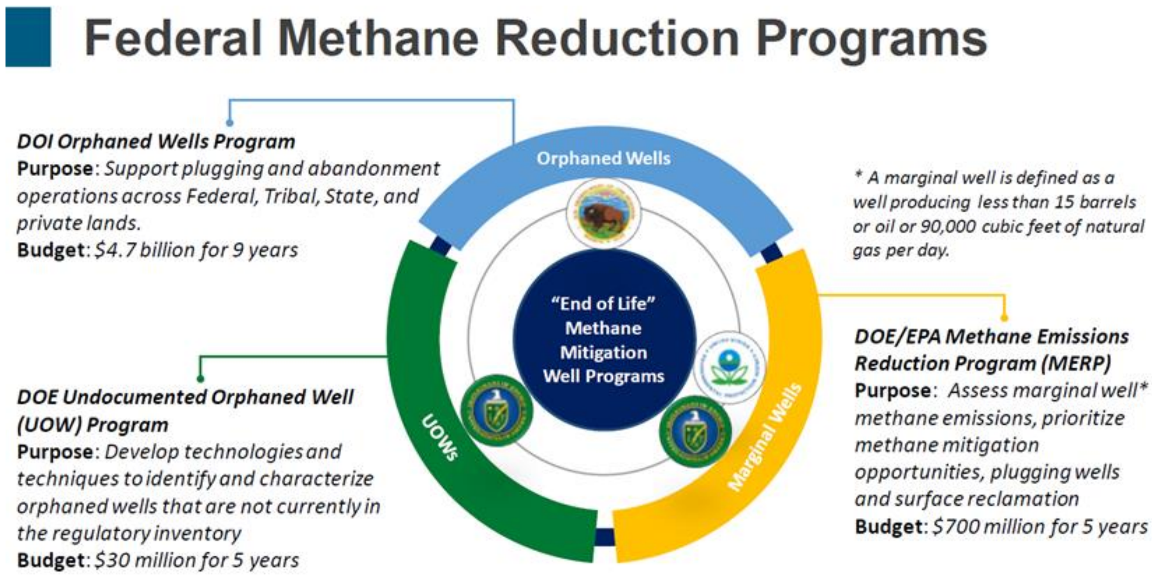


Figure 1.1: BIL funded programs targeting methane mitigation and reduction from historical oil & gas wells. Due to the nature of the challenge these bills address there are many points of overlap. Federal agencies are working to collaborate across these programs to provide for the most efficient use of funds.

UOWs are wells that lack an operator of record (orphan well), and do not exist in regulators' inventories. The estimated number of UOWs reported by states is between 210,000 and 746,000, while other sources estimate the number could surpass 1 million. As with orphan wells, UOWs provide uncontrolled pathways for methane and other substances to migrate upwards towards the surface where they can interact with groundwater and enter the atmosphere. The presence of UOWs is a major geotechnical hazard when developing subsurface resources (oil & gas, geothermal, carbon capture utilization and storage or CCUS, etc.).

Through CATALOG, UOWP will support the DOI Orphan Well Program Office (OWPO) and Methane Emissions Reduction Program (MERP) – a joint effort among the Environmental Protection Agency (EPA), NETL and the DOE's Office of Fossil Energy and Carbon Management's (FECM's) Division of Methane Mitigation, by providing technical and methodological assistance (Figure 1.2). The BIL provides investments to plug these wells through DOI OWPO, which will help communities reduce methane emissions and eliminate other environmental impacts. This investment is part of the Administration's overall response in remediating environmental concerns, addressing legacy pollution that harms communities, creating good-paying jobs, and advancing long overdue environmental justice.

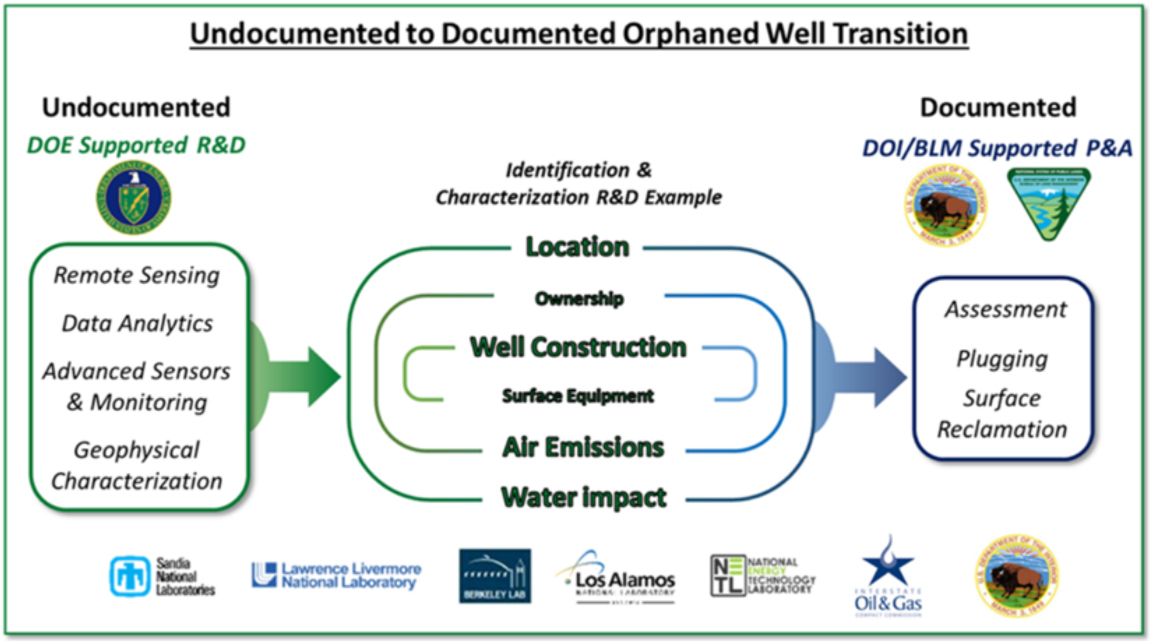


Figure 1.2: Demonstrating the flow of information from CATALOG to the DOI OWPO's $4.7 billion orphan well plugging program.

UOWP-CATALOG consists of 9 work packages that focus on topics and technologies related to identifying and characterizing UOWs. CATALOG is structured as an applied science program with emphasis on directly impacting field level activities of locating and characterizing UOWs. A summary of the key work packages' accomplishments towards these goals are detailed in this chapter. For a more detailed discussion of work package activities please see the relevant chapters (Chapters 3 through 11).

2. Summary of Key Work Packages

Work Package 1: Methane Detection and Quantification

WP1 Accomplishments focus on finding a simple, accurate and cost-effective method to quantify methane emissions from orphan and undocumented orphan wells, in order to prioritize their plugging through BIL funding. Current protocols to measure orphan well methane emissions require deploying highly trained personal with specialized equipment to measure the methane flux at a cost of roughly $2500/well. Given the number of wells with potential emissions to be characterized, CATALOG was asked by the Office of the White House and Department of Interior to determine if a simpler, accurate, more cost-effective method could be employed. If successful, this approach can be adopted as part of the DOI Methane Measurement Guidelines. CATALOG developed a relationship between methane concentration and flow rate as a function of windspeed. The approach uses a combination of Gaussian plume modeling, inverse modeling, and uncertainty quantification to estimate flowrate from concentration.

The following three points summarize the findings of this research effort:

1. The Gaussian plume method is viable under a specific range of atmospheric conditions and CATALOG's ongoing validation supports its use as a reportable methane measurement method.
2. The Gaussian plume method simplifies DOI methodology as it can be used for multiple well types.

3. The natural system temporal variability of methane emission rates from wells is within an order of magnitude relative to any single point measurement value.

CATALOG's Manvendra Dubey also contributed to DOI's report on "Assessing Methane Emissions from Orphaned Wells to Meet Reporting Requirements of the 2021 Infrastructure Investment and Jobs Act: Methane Measurement Guidelines", the July 2023 version:
https://www.doi.gov/sites/doi.gov/files/orphaned-wells-methane-measurement-guidelines-july-2023-version.pdf

Work Package 2: Well Identification

WP2 efforts during year 1 focused on an evaluation of geophysical techniques currently in practice for well identification, as well as techniques likely to provide useful, supplemental data streams to either minimize uncertainty in orphan well location identification or to add value in determining the physical characteristics of orphan wells. Coupled with the actual geophysical data to be collected is consideration of the acquisition method: ground survey, unmanned aerial vehicle (UAV) or helicopter. Each of these acquisition scenarios presents benefits (e.g., reduced survey time in densely vegetated areas) and drawbacks (e.g., higher cost) and there is no single scenario which is suitable for all likely field situations, thus requiring survey design to be done on a case-by-case basis. Similarly, WP2 examined two broad classes of geophysical data streams useful for well identification: magnetostatics and broadband electromagnetics. The former is comparatively mature and has been successfully employed through ground, UAV, and helicopter surveys for well identification, however the method is known to be insensitive to wells whose casings are either absent or constructed of non-ferrous materials. Broadband electromagnetics, from direct current (DC) to ground penetrating radar (GPR) frequencies, is shown to be useful in scenarios where magnetostatic methods fail, and because of the underlying physics, can provide information about well characterization such as wellbore integrity and fluid interactions with regional aquifers. However, an optimal survey design for electromagnetics remains an open operational question. Drone-based time-domain methods and GPR are presently in the field-testing stages. Scalable whole-field illumination via top-casing electromagnetic excitation is, in principle, a viable strategy for large scale survey execution and will be the subject of Year 2 efforts. Operational tradeoffs between magnetics and broadband electromagnetics are summarized in Chapter 4.

Lastly, extensive stakeholder engagement was conducted to develop best practices and to determine the current state of the art. It is clear the UAVs will play a critical role in well identification since in many cases they are expected to be more economical than either ground surveys or helicopters. The Work Package 2 team developed a list of "best practices" that describes the optimal flight parameters when deploying UAVs to locate UOWs. A key finding was that that nearly complete well inventories could be obtained if the flightline spacing was ≤ 50 m. In addition, nearly complete well inventories could be obtained by aeromagnetic surveys flown at altitudes ≤40 m.

Work Package 3: Sensor Fusion and Data Integration with Machine Learning

UAVs and machine learning are two technology spaces that have witnessed tremendous advancements in the last 5 years. WP3 is focused on combining these technologies to locate and characterize UOWs.

UAVs are capable of collecting multi-modal data that when processed by machine learning algorithms can detect UOWs more effectively than data from a single sensor, which is what is traditionally done today. The WP3 team has procured prototype sensors (e.g., magnetometers, methane sensors, lidar and ground penetrating radar sensors) and these are currently being integrated for installation on UAV platforms. Once ready, these UAVs will be used to collect the multi-modal data that are critical for advancing a robust method for locating UOWs.

Numerous data integration tasks are underway. The team is testing the utility of classical statistics and machine learning to predict well location based on various geospatial and remote sensing data types. The team has also identified gaps in the California state data bases and is working with the CALGEM state agency to obtain a full set of all well records in the state. CATALOG tools will be used to extract location and characterization information from these types of well records. CATALOG is also applying computer vision models to historical topographic maps and aerial photos to locate wells. Finally, CATALOG has released an integrated dataset consisting of private and public well data sets and selected attributes (age, type, depth, etc.) on EDX (https://edx.netl.doe.gov/dataset/co2-locate). This activity has been spun off into a new work package (Work Package 8) with the goal of tracking undocumented orphan wells that become documented orphan wells.

Work Package 4: Characterization

WP4 has focused on determining the physical characteristics of orphan wells that can be used to inform plugging and abandonment plans. The methods being developed and evaluated are: (1) Time Domain Reflectometry (TDF), which consists of propagating an electrical pulse down the well casing and recording various reflections, also recorded at the well casing, associated with damage or discontinuities, (2) Ground Penetrating Radar (GPR), which consists of airborne measurements that are sensitive to both cased and uncased wells, (3) Acoustic Characterization, which consists of recording traveling waves in the well casing and using travel times to determine the depth of the well, and (4) Direct Current Electrification (DCE), which consists of applying a current at the well casing and measuring the electrical field surrounding the well either at the surface or in a nearby well, revealing depth and casing damage information about the well.

Of these four methods, three require direct access to the well head or well casing, while the GPR method makes airborne measurements in the local vicinity. GPR can potentially detect unknown wells (a goal of the overall program, but not of this work package). Each method has some potential strengths and weaknesses, both in terms of what information they can provide, and the cost and complexity of the measurements themselves. The team is deliberately focusing on methods that do not require putting a tool or sensor into the well, as that adds cost and complexity. Ultimately, recommendations may be made for the application of simple down well techniques like lowering a camera or some other low-tech sensor, but these options are not currently being investigated. It is understood that a suite of measurements that can be performed simultaneously or synergistically, within this work package and with other work packages, will likely be the most effective way to characterize a well for plugging and abandonment.

Work Package 5: Integration and Best Practices

An efficient transfer of knowledge gained as part of the CATALOG program is needed to enable the transition of research & development results into practice. CATALOG's goal is to quickly develop best practices so that the DOI OWPO funding to plug and abandon wells can be used by states efficiently. Best Practices in the following areas were developed by CATALOG over the last year:

1. Working with WP1, WP5 Developed and is testing a methodology for quick and inexpensive methane emissions estimates from orphaned wells. The methodology utilizes downwind concentration measurements and wind field measurements, coupled with a plume dispersion model, to estimate methane emissions. Although the methodology does not provide the same high level of accuracy as flux chambers or high flow sampling, it allows for easy identification and classification of wells into one of three categories: (1) not leaking, (2) leaking at low or mid-level rates, or (3) high emitters that may warrant additional investigation. With this methodology, wells that should be assigned a high priority for plugging due to their relatively high methane emissions can be quickly and inexpensively identified. The procedure for utilizing this method is being documented for incorporation into a best practice publication.
2. The CATALOG team has reported a methodology for systematically compiling publicly available digital data prior to embarking on field campaigns. Geographical Information System (GIS) technology is used to analyze digital data and create maps to guide field work activities based on the goals of a specific project. This workflow was developed in the Appalachian region and although certain aspects may be unique, the general process should be applicable to locating undocumented wells in other regions. The report on the methodology was designed to assist states with UOWs, particularly those in the Appalachian Basin, in their initial searches for well locations across an identified area of interest.
3. The CATALOG team completed a report on best practices for conducting aerial magnetic surveys for locating UOWs. The sharing of the report will allow stakeholders to quickly benefit from the decade of experience that NETL has acquired in testing and demonstrating this particular technique.
4. For field testing and demonstration of technologies that are developed under the CATALOG program, the WP5 team has focused on: (1) assigning coverage of sites across the U.S. such that technologies and methodologies are relevant and successful in a variety of settings (i.e., different geology, topography, land use and vegetation cover, etc.), and (2) matching the needs of specific technology field tests and demonstrations with available sites.
5. WP5 has also created a glossary ("Key Terminology for the CATALOG Project") summarizing key relevant terms and associated generalized definitions, while noting certain synonymous state-specific terms, to help streamline discussions and ensure consistent understanding across all stakeholders.

Work Package 6: Data Management

The primary focus of the data management work package has been to develop the website for the CATALOG program, which can be found at: https://catalog.energy.gov/. The website provides an overview of the program as well as descriptions of the different work packages and access to the best practice manuals, CATALOG news and reports. In addition to the website, CATALOG's working documents have been placed on DOE’s SharePoint site for the internal team to access. In addition, DOE's virtual platform for public curation of R&D data and tools, Energy Data eXchange (EDX), is being used to share data among team members and stage data for eventual public release when appropriate.

FY23 Funded Work Packages

Work Packages 7-9 were stood up part way through 2023 and thus have only a partial year of work completed. Work Package 7 focuses on advancing Records Data Extraction, Work Package 8 is creating a Wells Data Base, and Work Package 9 focuses on the creation of Field Teams that will travel to different sites and apply the tools being developed in the other work packages. The overarching goal is to not only identify and characterize wells but to also work with State regulators to transfer the best practices being developed by CATALOG.

In October 2023, the multi-lab Work Package 9 team was deployed to Osage County, OK to locate and characterize orphan wells. The deployment employed a combination of CATALOG-generated data and tools including:

1. Estimates of orphan wells numbers and locations in the targeted area of Osage County collected prior to arriving at the site.
2. Drone flights of quad copter and fixed wing drones to locate wells using magnetometers.
3. Methane measurements for detection of emissions from leaky wells using both ground-based and drone investigative teams.

The deployment was highly successful as a first attempt with many wells being identified and several large leaks identified for plugging. The team is currently processing the data from this deployment. Osage officials and the Work Package 9 team are planning future trips to the area for further investigations.

3. Work Package 1- Methane Detection and Quantification

The following team members contributed to this chapter: Sebastien Biraud (LBNL), Manvendra Dubey (LANL), Mohit Dubey (LBNL), Emily Follansbee (LANL), Raghava Gorantla (NETL), James Lee (LANL), Natalie Pekney (NETL).

3.1 Section Summary

Work Package 1 is focused on methane detection and quantification, which had three major priorities identified for FY23:

1. Development of a baseline evaluation of existing methane emission quantification technologies, approaches, and promising new techniques,
2. Development of advanced methane quantification technologies and techniques that meet the needs of the States and Federal agencies responsible for Federal and Tribal Lands, and
3. Field demonstration of efficient, accurate and cost-effective technologies and methods developed in this work package.

The multi-lab WP1 team is working together and sharing knowledge and resources on all three priorities described above. Their combined expertise for detection and quantification of methane point sources within the energy sector has led to several quick research “wins.”

Priority 1 is an on-going effort, as methane detection and quantification technologies are evolving and new technologies are emerging all the time. This activity overlapped with WP3 (Sensor Fusion and Data Integration with Machine Learning) with the addition of methane sensing capabilities to a multi-sensor drone platform. One lab (LBNL) worked closely with the WP3 team at LLNL, hosting a visit and loaning a methane sensor (Aeris) for several months to the WP3 team for integration and assessment.

Engagement with multiple stakeholders at the Federal, state, and local level led the WP1 team to conclude that the current high cost of instruments and high level-of-effort needed to conduct a pre- and post-plugging methane emissions estimate were not scalable and did not meet current needs. One lab (LANL) is leading a new approach to quantify point source emissions from abandoned wellheads using a dispersion model, ambient methane concentrations and wind observations. This work led to a white paper that was shared with the CATALOG project Program Manager and posted on the project web site in January 2023. One lab (LBNL) is leading the laboratory and field validation effort of this promising new technique. This activity overlaps with the WP5 (best practices) effort to develop and test methodologies for a quick and inexpensive methane emissions estimate from orphaned wells.

Under Priority 3, several field deployments were conducted in FY23 to assess multiple methane detection and quantification technologies. This work is an on-going effort that will continue for the duration of the project in coordination with WP9 (Field Team). In March 2023, NETL hosted a team from LBNL at Hillman State Park, PA along with the Pennsylvania Department of Environmental Protection, a Non-profit Organization (Adventure Scientists), and a manufacturer (Heath Consultants). In April 2023, the Well Done Foundation (https://welldonefoundation.org/) invited WP1 staff from LBNL and LANL to participate to an “Oil & Gas Well Field Camp” in Hobbs, NM. During both field deployments several methane sensing technologies were evaluated. Valuable observations were also collected to validate the WP1 low-cost approach to methane emissions estimation (Follansbee *et al.*, 2023 – in prep). In May 2023, the LBNL team led a field campaign near the Lost Hills oil & gas field in California that overlapped with WP3 (Sensor Fusion and Data Integration with Machine Learning) to validate a methodology for the detection of potential UOWs using historical topographic maps and aerial photos.

3.2 Activity

Activities carried out over the past year related to each of the three priorities listed earlier are detailed below.

Development of a baseline evaluation of existing methane leakage detection and emission quantification technologies, approaches and promising new techniques.
In the last year, WP1 teams procured, shared and tested methane sensors in a laboratory setting (for independent verification of the system performance claimed by manufacturers) and during field deployments (for evaluation of useability and performance in a real world environment). This work includes evaluations of the following sensors:

- Picarro Gas Scooter;
- AERIS Strato;
- SEMTECH Hi-Flow2 and RMLD;
- Chambers-based systems (static and dynamic);

- OGI cameras;

One lab (LBNL) also engaged with a startup (XplotRobot Inc, https://xplorobot.com/) to assess their integrated methane mapping sensor package (hand-held) for precise localization and quantification of methane emissions from all types of oil and natural gas handling equipment. This system enables localization of the point sources to the specific part of the equipment so the field operator (Figure 3.1) can conduct timely mitigation measures. Figure 3.1 shows an application for a natural gas compressor station.

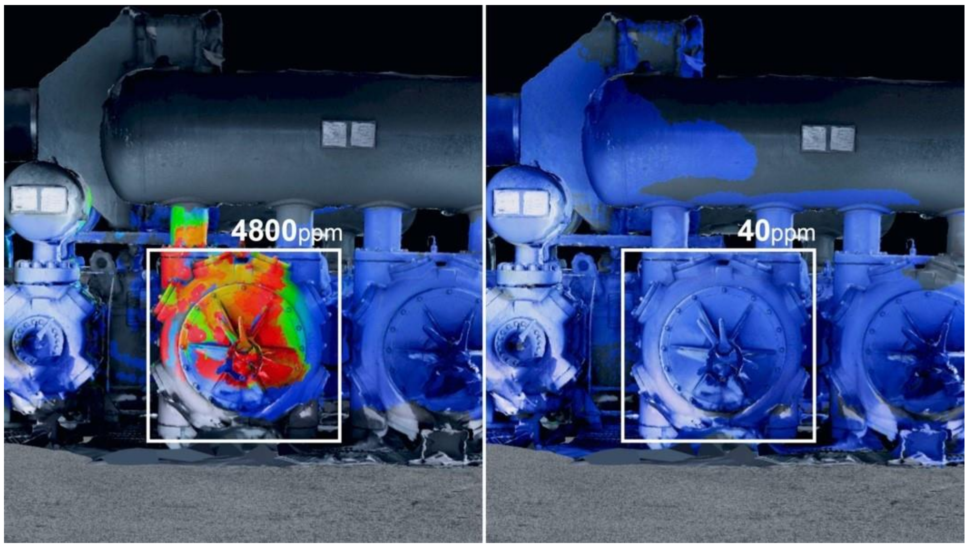


Figure 3.1: Example of equipment assessment –methane concentration mapping and emission source identification on a Digital Twin of a compressor inside an operating compressor station. Left: leak identification, Right: post-repair confirmation of emissions mitigation.

Development of advanced methane quantification technologies and techniques that meet the needs of the States and Federal Agencies responsible for Federal and Tribal lands.

In the last year, the WP1 team engaged with stakeholders during several in-person, hybrid and remote meetings and field deployments:

- Organized a meeting with Energy-ICorps conducting customer discovery/needs engagement for LDAR-Q (e.g., Champion-X, MIX, Duke Energy, CVX, GTI, CVX, Pioneer);
- Participated in a "CH4 Connections" GTI conference in Ft. Collins;
- Attended a AAPG workshop on UOW in Oklahoma City, OK;
- Participated in discussions with Ted Boettner, Ohio River Valley Institute, and two members of the Texas Railroad Commission about how to measure methane emissions from orphaned wells;
- Hosted a briefing to the CATALOG program by Doug Baer (Los Gatos/ABB, http://www.lgrinc.com/) on the use of laser sensors to measure atmospheric CH4
- Hosted a briefing by Curtis Shuck (WellDone Foundation, https://welldonefoundation.org) and explored collaboration on assessing hardware/best practices (operations and costs) for leak quantification pre- and post- orphan well plugging;
- Was an Invited Speaker at The Capitol Forum Energy and the Regulated State Event - Natalie Pekney (NETL), "Infrastructure Investment & Jobs Act In Action, National Inventory & Tracking Unknown Wells";
- With KY Geologic Survey, developing potential collaboration for measuring methane emissions from orphaned wells in KY on private properties;

- With KY/U.S. Forest Service, planning collaborative research on wells in the Daniel Boone National Forest;
- With NY Department of Environmental Conservation, carried out multiple discussions regarding methane emissions measurement techniques and results from field campaigns NETL has completed in Olean, NY;
- Carried out field observations in Hobbs, NM monitoring orphan leaks with WellDone Foundation;
- Carried out field observations at Hillman Park, PA with community stakeholders including Adventure Scientists;
- Carried out field observations in McKittrick, CA in coordination with WellDone Foundation;
- Carried out a field scoping trip in the Four Corners area (CO, UT, AZ, NM) with local stakeholder group and tribal communities;
- Maintained active engagement with DOI – MIT, meeting biweekly on orphan well screening protocols;
- Partnered with New Mexico Tech - Socorro to deploy CH4 sensor on UAV in the field.

These discussions with stakeholders led the WP1 team to conclude that a novel, simple, defensible, safe, and cost-effective method to quantify methane emissions from undocumented abandoned orphan wells was needed. Our proposed method was to measure excess methane concentrations over background at multiple points downwind of a leaky well together with wind speed and direction to infer the emission rates using a Gaussian plume model. It is worth noting that gas observations downwind of an *unknown* emission point source location are routinely used to infer emission rates and calibrated with Gaussian models for compliance with air-quality regulations (e.g., EPA's OTM 33A). Our method takes advantage of the known location of the emission point source and the relatively short distance between the point source and the methane and wind sensors. This method is described in a white paper posted on the CATALOG website (https://catalog.energy.gov/publications/). In this paper, we demonstrate the potential value of this approach to estimate emission rates from orphan wells in controlled laboratory releases at LBNL for flow between 1 to 30 g/hr. Figure 3.2 shows a schematic of the experimental setup.

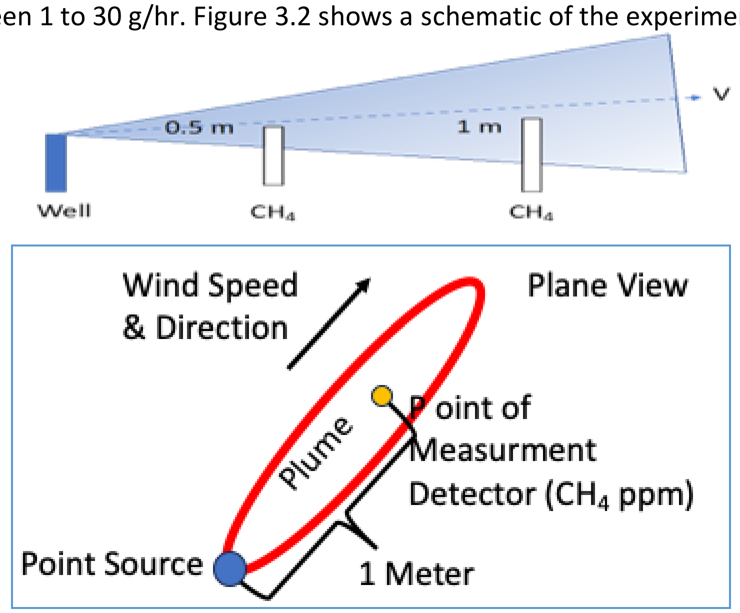


Figure 3.2: Experimental setup for our plume monitoring and gaussian plume model analysis to infer emissions based on wind statistics and methane concentrations measured downwind of the point source.

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 simple solution that we have tested in laboratory conditions that can be easily deployed in the field. 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 to assess point-source strength. 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 that underpins the principle of our method (Figure 3.3).

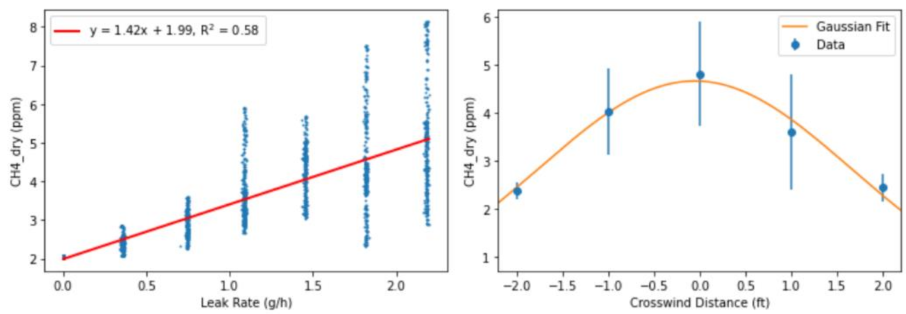


Figure 3.3: (Left panel) Measurement of downwind methane concentration (1 m) from the point source, aligned with the center line of the plume as a function of the leak rate. The red line is a fit through the measured methane concentrations. Methane concentration variability is caused by the change in the position of the plume as shown in the Gaussian plume profile (right panel).

The proposed method involves measuring methane concentrations at a few known distances downwind from the point source for a few minutes in order to capture the plume structure. The crosswind and downwind profiles can then be combined with dispersion models to estimate the well's emission strength. The accuracy and threshold of our emission estimate, which depends on the sensitivity of the instrumentation gas analyzer used to measure methane concentrations, spans a wide range from 1 to 100s g/hr. Particularly sensitive measurement techniques (at the ppm level) allow for cost effective determination of smaller leak rates (<10 g/hr) that current methods are unable to quantify and that continue to be a challenge. This issue is particularly critical for the UOW problem, given the large number of these wells, whose overall methane emissions, despite the small leak rate per well, can add up to a large contribution to overall emissions.

The WP1 team tested this technique, together with the WellDone foundation, on high emitting (0.5 to 10 kg/hour) orphan wells in Hobbs, NM that have been prioritized for plugging. We demonstrated that our ambient methane and wind measurements, when used in Gaussian-plume inversion, have the potential to derive emission rates with accuracy sufficient for prioritization (better than a factor of two). Our results also showed that emissions from leaking wells that have been shut-off for months, can increase by a factor of two when re-opened, due to pressure buildup. This systematic high bias needs to be corrected for when reporting on the methane emission reductions resulting from plugging.

Field demonstration of efficient, accurate and cost-effective technologies and methods developed in this work package.

During field deployment, WP1 National Laboratory partners worked together to leverage access to areas where a variety of oil and natural gas wells are present (i.e., undocumented orphaned, orphaned, idle or shut in, and producing). These field locations represent unique opportunities to test and assess methane detection and emission quantification methods. Securing access to such facilities is critical to the WP1 mission but can be very challenging and time consuming.

Below are a few examples of deployments in FY23:

- *Deployment at Hillman State Park, PA*: Natalie Pekney (NETL) led several field campaigns to this area. In March 2023, a team from NETL and LBNL worked together to measure gas concentration and composition from orphan wells prioritized for plugging. A Picarro backpack and RMLD sensor were deployed to detect methane leaks, a FLIR OGI camera was used to find leakage points, and a Xplorobot LIDAR and SEMTEC HI-FLOW2 sensors were used to quantify methane leak rates at the well head. It is worth noting that leak rates at this location range between 10 and 100 g/hr, which is relatively small. Our sensitive sensor measurements detected methane concentrations of a few ppb around the surface of well locations. This illustrates how sensitive sensors can indeed find small leaks, and the coordinated use of a FLIR camera and a SEMTECH HI-FLOW2 sensor can enable methane quantification (but this is not low-cost).

Figure 3.4: Gas concentrations and composition, and flow rates measured at orphan wells in Hillman State Park, PA. Upper left: Xplorobot sensor; Upper right: RMLD sensor; Lower left: staff from PA-DEP collecting a sample for gas composition analysis; Lower right: NETL and LBNL Team.

- *Deployment at Hobbs, NM*: The WellDone Foundation, a non-profit organization that has been contracted by several states to plug orphan wells across the USA, invited members of the WP1 team to participate in a field campaign in April 2023. An identified high emitting orphan well was targeted for observations on 19 April 2023. The well is located southeast of the city of Hobbs, New Mexico and was at pre-plug status during the time of sampling (Figure 3.5). It had been drilled to 6500' depth with historical production from the San Andres and Blinebry formations. We visited the well site during a planned pressure release. Due to pressure buildup of natural gas within the well and uncertain integrity of the wellhead and casing infrastructure, some orphaned wells are periodically vented to relieve pressure. This safety related venting provided an opportunity to test our approach, to sample downwind of a well during venting operations at known release rates, and to carry out prior-to-plugging measurement of typical fugitive emission rates under similar atmospheric conditions.

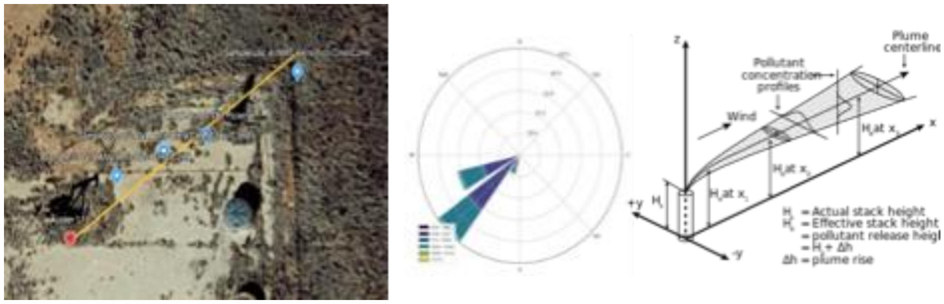


Figure 3.5: Field set up, wind data and plume model used in our Hobbs, NM orphan well campaign in April 2023.

The wind direction and methane observations were collected at 7.5 m, 15 m, 22.5 m, and 47 m downwind of the source are shown in Figure 3.6. All data are resampled to 1-minute averages. The production tube valve was opened at 11:30 am to vent, creating a large methane point source. While the well was vented, large peaks of up to 50 ppm methane concentration in excess of pre-venting methane concentrations were observed at 7.5 m, 20 ppm excess at 15 m, 8 ppm at 22.5 m, and 3 ppm at 47 m. Observed methane signals at different sample locations are highly correlated with larger concentration enhancements at sites closer to the sources for both the high and low leaks. Mean enhancements during venting were 37.44, 18.01, 7.91, and 2.76 ppm at 7.5, 15, 22.5, and 47 m respectively.

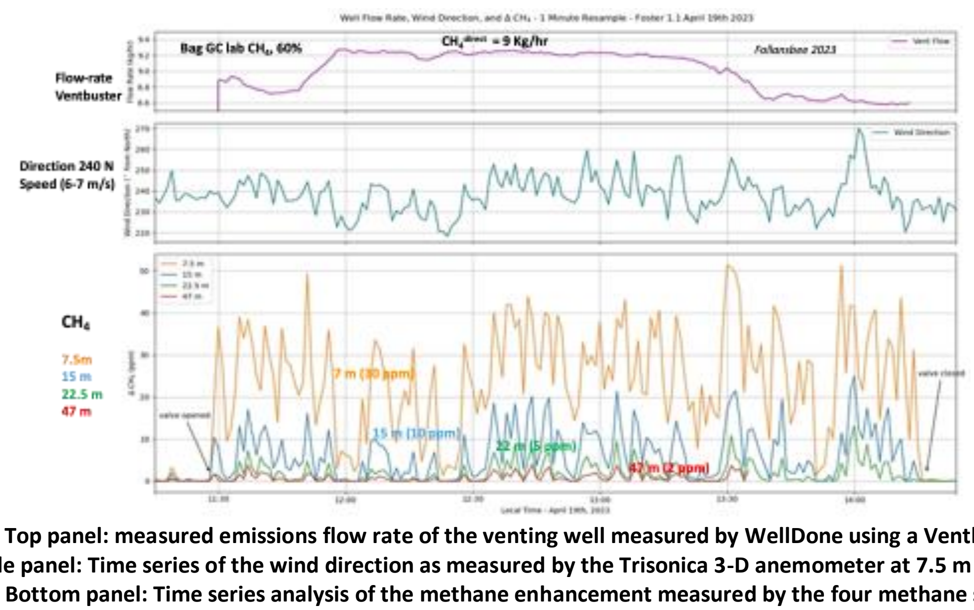


Figure 3.6: Top panel: measured emissions flow rate of the venting well measured by WellDone using a Ventbuster flow meter. Middle panel: Time series of the wind direction as measured by the Trisonica 3-D anemometer at 7.5 m downwind of the well site. Bottom panel: Time series analysis of the methane enhancement measured by the four methane spectroscopic instruments downwind of the venting Foster 1.1S well at 7.5 m, 15 m, 22.5 m, and 47 m. Annotations mark when Foster 1.1S vent was opened and when the vent was closed.

Observations collected during this field deployment and results from our approach described above are being included in a publication that will be submitted to a scientific journal for peer-review before the AGU Fall Meeting in December 2023 (Follansbee *et al*, in prep).

- Finally, one National Lab partner (NETL) is examining the temporal variability of emissions from orphan wells using Earthview BluBird sensors/Flux Chambers (Figure 3.7). This work is being done at several locations in the US northeast (PA, OH, NY). Preliminary data show a factor of 10 variation in sub-daily leak rates that are lower than standard High Flow Sampler snapshot data, indicating that short-term sampling can be used for prioritizing wells for plugging and abandonment. The team is continuing active monitoring to better understand the cause of the emission variability.

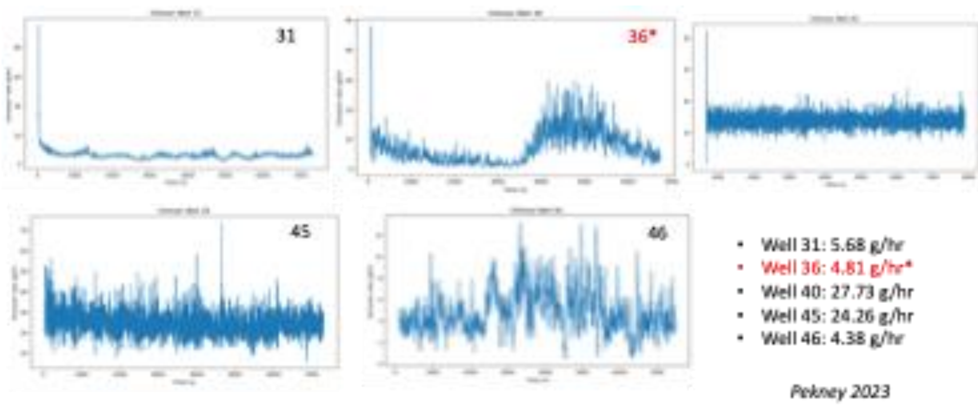


Figure 3.7: Example of short-term variability in methane emissions from several wells in Hillman State Park, PA.

3.3 Motivation and Background

Orphaned and abandoned oil and gas (O&G) wells contribute significantly to anthropogenic methane sources in North America and to global climate change. There are an estimated 2.3 million abandoned and orphaned oil and gas wells in the United States emitting approximately 200 Gg CH_4 per year. In November 2021, the United States Congress passed the Bipartisan Infrastructure Law (BIL) with sections aimed at screening orphaned O&G wells for prioritized plugging and abandonment (P&A). Nationwide, P&A operations are challenging with current methods of locating and identifying orphaned O&G wells prohibitively costly and time consuming. Faster and cheaper technologies for locating orphaned and abandoned wells and prioritizing them for well plugging are needed.

The nationwide ecosystem of UOWs is very diverse, with heterogeneous surface and subsurface conditions across oil and gas basins. Determination of methane emissions from these wells is key to quantifying greenhouse gas impacts of UOWs and to evaluating the benefits of successful well plugging. We currently have no established methodology for making methane emission measurements in different geological and environmental settings as a function of scale.

The objectives of WP1 are to advance existing state-of-the-art methane detection and quantification technologies and approaches and/or develop new technologies designed to cost effectively and efficiently find (identify) and assess (characterize) methane emissions from UOWs so that they can be prioritized by Federal, State and Tribal Agencies (Stakeholders) for mitigation. Given the significance of accurate methane emissions measurements to demonstrating successful implementation of this program, methane detection and quantification has its own work package but makes direct contributions to the identification (WP2), sensor fusion and machine learning (WP3), characterization (WP4), and best practices (WP5) work packages.

3.4 Data Summary

All of the collected field data for controlled releases on orphan wells in Hobbs, NM, Hillman State Park, PA and from the controlled release campaigns summarized below, are publicly available.

- Permian basin: Hobbs, NM
  - ***HL2 #3: 32.68855N, -104.05294E***
    - Estimated flow rate from WellDone **147 g** (CH4)/hr, composition 60% CH4 (m/m), production depth 3150 ft
  - ***HL2 #1: 32.69244N, -104.05294E***
    - Small leak, no documentation
  - ***Foster 1S: 32.69125N, -103.07491E***
    - Estimated flow rate **9 Kg/hr**, composition 60% CH4, production depth 3700 ft
- San Juan basin: Farmington, NM
  - Visited four wells on BLM, one was confirmed small leak
  - NE Hogback #5 Unit C: 36.820362N, -108.517998E, Production depth 1537 ft
- Hillman National Park, PA
- Controlled Releases LBNL, CA
- Controlled Releases LANL, NMT, Socorro, NM

Our results are being reported and shared in two peer-reviewed manuscripts (in preparation) for AGU's 2023 fall meeting, and key findings are being regularly shared with the BIL's interagency MMT and IOGCC to maximize their impact on reducing methane emissions on UAOWS.

Key Findings:

1. Gaussian plume method is viable and our ongoing validation supports its use as a reportable methane measurement method.
2. Gaussian plume method simplifies DOI methodology as it can be used for multiple well types.
3. The natural system temporal variability of methane emission rates from wells is within an order of magnitude relative to any single point measurement value.

3.5 Future Work

The WP1 team is planning field campaigns with WP9 in Osage County, OK and in the Four Corners areas to evaluate our dispersion model technique. Furthermore, the team is evaluating the sensor sensitivity and speed needed in order to reduce costs and ease operations for emission quantification. The WP1 Team is working with WP9 on deploying a methane sensor on a fixed wing UAV together with a magnetometer and will use wide area surveys to find undocumented wells at lower costs. The team has focused on wells that are single point sources and that are accessible and plans to extend the approach to wells with infrastructure that makes access challenging. In addition, the team will examine the smaller diffusive leaks from sub-surface pathways around the wells using steady state chambers in collaboration with the characterization efforts of WP3. The WP1 Team is also providing technical inputs on protocols and best practices for operators and states with other work packages.

4. Work Package 2- Well Identification

Contributors to this chapter are: CJ Weiss (WP lead, SNL), R Hammack (WP co-lead, NETL), Y Wu (LBNL), GD Beskardes (SNL), S Mukherjee (LLNL) and C Morency (LLNL).

Exploration and development of oil and gas resources across the multiple geologic provinces of the continental United States over the last 150 years has resulted in a deep historical record that reflects the economic vibrancy of the industry, the strongly heterogeneous approaches to regional oversight, and the evolving technologies available for documentation. As such, a complete accounting for all oil and gas wells drilled in the US is both unconsolidated and inherently incomplete, as their presence has been obscured through decades of compounded overprinting by such dynamical processes as population encroachment, land-use diversification, natural erosion and vegetative growth. **Our goal is to develop geophysical sensing technologies to fill those gaps in the well inventory.** The challenge is rooted in both the overwhelming number of unidentified wells, and the necessity of creating new methods for merging disparate data sources to efficiently map their location.

By developing geophysical sensing technologies to identify undocumented orphan wells (UOWs), **the WP2 team informs prioritization (WP5)** by Federal, State and Tribal Agencies (Stakeholders), follow-on **characterization (WP4)**, **data integration (WP3)** and **field demonstrations (WP9)**. To be relevant, these technologies and approaches must be scalable to the challenge of identifying 10s of thousands of wells per year over a complex and diverse set of operational constraints and well designs. Our objective is to meet this demand by maximizing workforce engagement in a three-pronged approach: layperson solutions (e.g. smartphone apps) accessible to tens of thousands – or more! – of non-specialist "citizen scientists"; commodity multi-physics systems for a smaller set of trained geotechnical professionals looking at more challenging identification scenarios; and, for the most difficult cases, specialized hardware/software systems requiring advanced (e.g. PhD) analytical expertise.

4.1 Section Summary

Under WP2 there are three principal work priorities.

1. Engage with stakeholders and geophysical contractors to assess field sites, operational considerations, and information gaps in the identification problem.
2. Develop scalable identification technologies, both in hardware and analysis, to meet the demands of geophysical sensing of 100's of thousands of wells across variable terrain and environmental conditions.
3. Develop identification strategies, both in hardware and analysis, to find wells in "hard" environments with features such as rugged topography, urban encroachment, dense vegetative cover, etc.

The WP2 team (SNL, NETL, LLNL and LBNL) has addressed these priorities in a number of ways. Primary among them is an initial breakdown of considerations in site characteristics and survey designs (Table 4.1) which affects what type of survey can or should be deployed for geophysical sensing. We restrict ourselves to geophysical surveying methods based on magnetometry and electromagnetic sensing as these are amenable to both on-the-ground, and airborne data acquisition scenarios (Table 4.2). This addresses **priority 1** listed above.

Site/Survey Factors	Ground Survey	UAV Survey	Helicopter Survey
Survey size	< 10 acres	10 – 1200 acres	>1200 acres
Visual Line of Sight	N/A	500 m max, no obstructions	Limited only by atmospheric visibility
Site Accessibility	Road, marine or aircraft	Road, marine or aircraft	N/A
Percent Forest Cover	0 – 100%	0 – 80%	0 – 100%
Terrain	Almost all	Moderate terrain	Moderate terrain
Proximity to Airports	N/A	Airspace restrictions	Airspace restrictions
Cost	Low	Moderate	High
Speed	Slow	Moderate	Fast

Table 4.1: Selection criteria for ground, UAV and helicopter surveys for magnetic and electromagnetic data collection.

	Magnetics	Electromagnetics (DC to GPR)
Physical property mapped	Magnetic susceptibility (e.g. high for iron, magnesium)	Electrical conductivity and permittivity (for metals, brines, clays)
Sensor size	mm to cm	~1 m or greater
Ground and/or Airborne?	Both	DC – ground only, otherwise both
Active or passive sensing	Passive	Both
Source required?	No	Both
Application	Location of steel cased wells	Location of both cased and uncased wells **(WP2)**, subsurface characterization **(WP4)**
Utility for well characterization (WP4)	Limited	High
Concerns/opportunities	Signal overprinting in cluttered environments; opportunities for gradiometry with drone swarms	Ambient EM noise and clutter noise; opportunities to exploit infrastructure to illuminate large field areas
Data Reduction Effort	Low to moderate – mag and flag isolated anomalies with corrections for drift, positioning, etc.	Low to high – target anomalies can be complex, requiring advanced numerical modeling to interpret
Where is CATALOG testing it?	OK, PA, WV	CA, OK
Deployment platform	Quad and fixed wing UAV	Quad UAV

Table 4.2: Selection criteria for magnetic and electromagnetic data collection

The WP2 Team (SNL, NETL, LLNL and LBNL) has also focused its efforts toward end-users on the more technically trained (2/3) side of the usability spectrum described above, **addressing priorities 2 and 3.**

Layperson engagement through smartphone app development is redirected going forward in the project as a separate **WP10** (LBNL Lead, SNL Co-lead) described later in this report. Specifically, the WP2 team has leveraged the existing expertise in UAV-based magnetometry (NETL) in two key areas: development of supplementary geophysical data collects and analyses based on ground/well electromagnetic responses (SNL, LBNL and LLNL); and establishment of a best-practices rubric for UAV operations (NETL). Focusing on UAV sensing platforms for near-ground reconnaissance is the operational scenario mostly likely to meet the scale of identification task due to the relatively short data collection times (roughly 320 acres/day for a single drone), the fact that they can be run in parallel over multiple survey sites, and the rapid development of this particular technology space (e.g. self-guidance systems, swarms, etc.). The focus on electromagnetic (EM) methods is based on their compatibility with UAV operations in a stand-off sensing posture where data can be continuously collected at altitude without direct coupling between the EM sensor and the ground. Such methods are also potentially sensitive to wells where the casing has been removed or is no longer intact. It should be noted that many of the sensing and analysis methods developed in WP2 have dual use in the characterization task **WP4** and there is overlap in personnel between these two tasks to best facilitate impact in each. Lastly, note that the computational burden of analyzing electromagnetic data can be high, and that practical methods for reducing it are required if data is to be collected at scale. To meet this need, SNL has advanced its scalable HiFEM computational library to include geophysical electromagnetics, thus enabling experiment design testing, field data reduction, site/well characterization **(WP4)** and an engine to inform ML algorithms in data integration **(WP3).**

4.2 Activity

4.2.1 Magnetometry (Priorities 1-3, NETL lead)

Stakeholder Engagement (Priority 1): The WP2 team has actively sought out stakeholder input for best practices, awareness of the state of the art, field site access and collaboration in technology development. Determination of acceptable flight parameters for aeromagnetic surveys was informed by input from CGG Airborne Surveys (now Xcalibur Airborne Surveys) and ExxonMobil. Designing Aeromagnetic Survey for Stonewall Jackson Reservoir, West Virginia could not have been possible without cooperation from Bureau of Land Management, US Army Corps of Engineers, DOE-Office of Aviation Services, Juniper Unmanned, UAV Exploration and Skyfront. Finding study areas with uncased wells was done in coordination with DiGioia Gray, Consol Energy, EQT and Range Resources-Appalachia. Lastly, progress in developing ML methods to compensate aeromagnetic surveys for aircraft maneuvering benefitted from thoughtful discussions with colleagues from Colorado School of Mines, DiGioia Gray, Juniper Unmanned, Air Data Solutions and UAV Exploration.

Determination of Optimal Flight Parameters (Priorities 1-2): Between 2005 and 2014, NETL contracted helicopter magnetic surveys of two oilfields in Wyoming and six large historic oil and gas producing areas owned by the Commonwealth of Pennsylvania. Flightlines ranged from 25 to 30 m apart and flight altitude was nominally 20-30 m AGL. The tight flightline spacing and low altitudes were intentionally conservative because the objective was to develop a complete inventory of existing wells. However, the tight flightline spacing resulted in greater survey cost (See Table 4.1); the low-altitude flight brought the aircraft closer to terrain, trees, and powerlines. The working hypothesis was that these surveys were oversampled and that an accurate well inventory could be obtained more safely and with less cost using wider flightline spacing and higher altitudes. This hypothesis was tested by removing flightlines of data from two aeromagnetic datasets (one in Wyoming and one in Pennsylvania) to simulate surveys flown with wider flightline spacing. The remaining data were re-gridded, and a peak picker algorithm was run to provide an unbiased count of well-type magnetic anomalies. **The result of this simulation was that nearly**

complete well inventories could be obtained if the flightline spacing was ≤ 50 m. This result was consistent using either the Wyoming survey or the Pennsylvania survey. However, the locational accuracy is degraded with wider flightline spacing because gridding will shift the well anomaly to magnetic maximum on the nearest flightline i.e., wider flightline spacing results in greater well location uncertainty.

Aeromagnetic flights at increasing altitudes were simulated by calculating the upward continuation of the magnetic field for each reading in the original survey and then re-gridding the upward continued data. **This simulation determined that nearly complete well inventories could be obtained by aeromagnetic surveys flown at altitudes ≤40 m AGL.** However, this is not always practical because obstacle height (e.g., trees and powerlines) sometimes exceed 40 m.

Corrections for Aircraft Proximity and Maneuvering (Priority 1): The aircraft is a significant source of magnetic noise in aeromagnetic surveys. Although physics-based compensation systems exist to correct for aircraft maneuvering, they are too expensive in terms of needed hardware/software and processing costs to be used by the small drone operators who are doing most well location surveys. Commercially available magnetic sensors designed for use on UAVs are now recording compass, gyro, and accelerometer data in addition to the typical magnetic data and GPS data strings. **This provides an opportunity to use machine learning (ML) algorithms as a low-cost method to compensate for sensor orientation/attitude and the time-varying spatial relationships between the aircraft and sensor.** ML algorithms are being developed by the WP2 team that will be trained and tested first with synthetic well-type magnetic data and later with real data from the Stonewall Jackson Reservoir surveys.

Uncased Wells (Priority 3): NETL reached out to industry partners (Range Resources-Appalachia, EQT, and Consol Energy) to locate a potential field test area that contained uncased wells. Although these companies locate thousands of abandoned wells yearly as part of their normal operations, they could not identify with certainty, a location with an uncased well. However, a collaboration with Andrew Zorn (DiGioia Gray) identified two uncased wells using **historic air photos** and a drone magnetic survey to confirmed that there was no associated magnetic anomaly, thus demonstrating the importance integrating geophysical data with complementary data sources **(WP3).**

4.2.2 Time- Domain and Induced Polarization Electromagnetics (Priority 2-3, LBNL lead)

The focus of this effort is the development of a drone based transient electromagnetic system for UOW to provide complementary data to that collected by magnetometry for the case of steel cased wells, and to fill a sensitivity gap in magnetometry data where the casing is either absent or non-metallic. Analysis of the time domain data is approached from two directions. The first is its interpretation as an indicator of general electrical conductivity anomalies present in the Earth/well system (e.g. steel, sand, rock). The second is interpretation based on the less common, electric polarizability of materials commonly associated with drilling and production activity (e.g. clay rich pads or spill/contamination induced metal sulfides or magnetite precipitations) which manifest as an "induced polarization" (IP) TEM response.

Drone TEM system hardware: A modern drone-based time-domain EM (TEM) system was acquired in the past year and tested for its functionality of well finding. Figure 4.1 shows an example dataset where a metallic wellbore with known location was used to test the performance of the system. The TEM system is integrated with the drone in terms of flight path planning and data streaming. A coaxial transmitter/receiver antenna loop is used for the EM system, along with a bucking coil for noise reduction. An EM anomaly associated with the wellbore is clearly identified. **Tests suggest that the drone system should be limited to low flight (<3-4m AGL) to maintain sensitivity to metallic objects.** Next steps are to modify the system design to allow greater flight heights (10 – 40 m AGL) while maintaining similar sensitivity. If successful, this will fill an operational gap between low flying drones (ill-suited for densely vegetated areas) and large higher-altitude, helicopter-based EM systems which have demonstrated limited sensitivity to the presence of steel-cased wells.

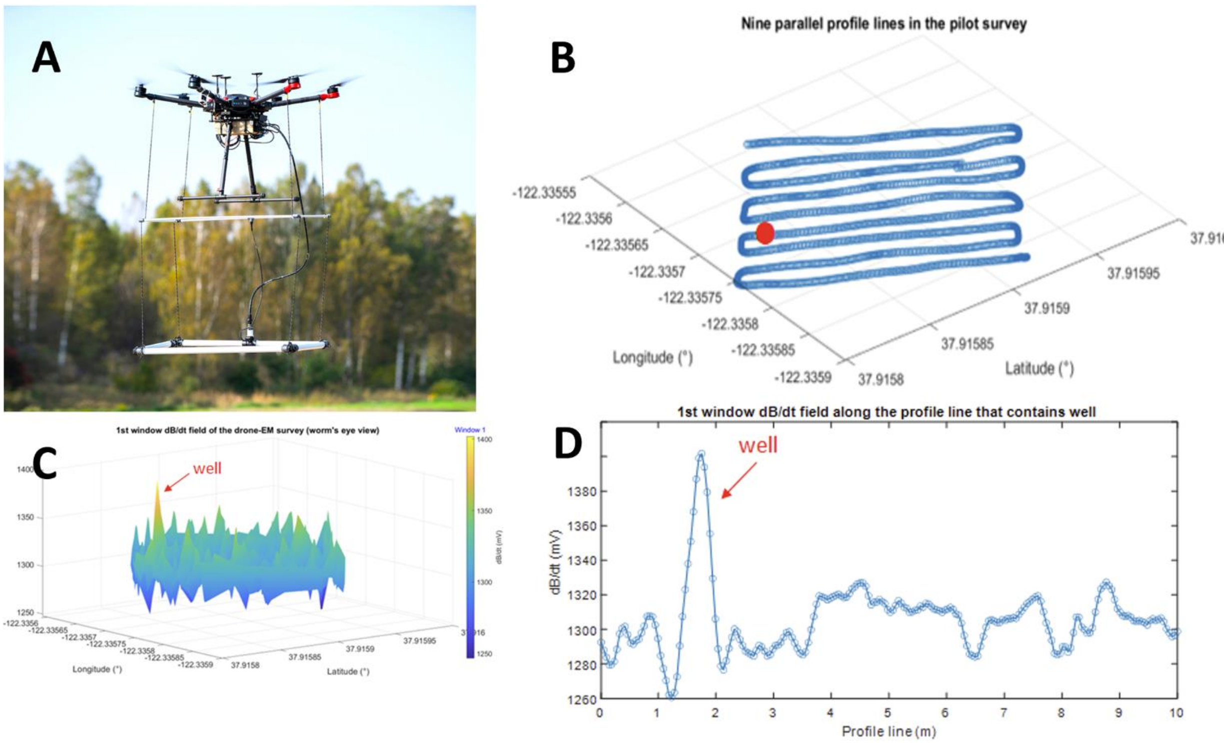


Figure 4.1: Drone based EM system and example dataset for steel casing identification. (A) Drone based EM system with autonomous flight planning; (B) flight path for test flight with red dot showing the location of the well; (C) EM anomaly identified from the EM data that is associated with the well; (D) A linear view of the EM signal associated with the wellbore in comparison with the background.

<u>1D Electromagnetic modeling for expected well response:</u> The operational principle of the TEM method is paused data collects (typically only a few seconds), where the Earth response from repeat transmitter wavelets is recorded and stacked to reduce instrument and environmental noise. Thus, the signal of interest is exponentially decaying magnetic field resulting from inductive excitation of the Earth/well system by the transmitter antenna. A "first cut" estimate of signal character and amplitude can be rapidly computed by a simple layered model where the presence of the well is approximated by its volume-averaged conductivity with the surrounding geology. Clearly this is a greatly simplifying assumption, but it at least gives a starting point for additional high-resolution modeling analyses. As an example, simulated drone TEM responses at 40 flight altitude show that signal levels for well (layered Earth) response are typically above expected noise floors and capable of distinguishing metallic from non—metallic (clayey) wells (Figure 4.2). Note that highly polarizable clayey wells generate a distinctive TEM response diagnostic of their presence.

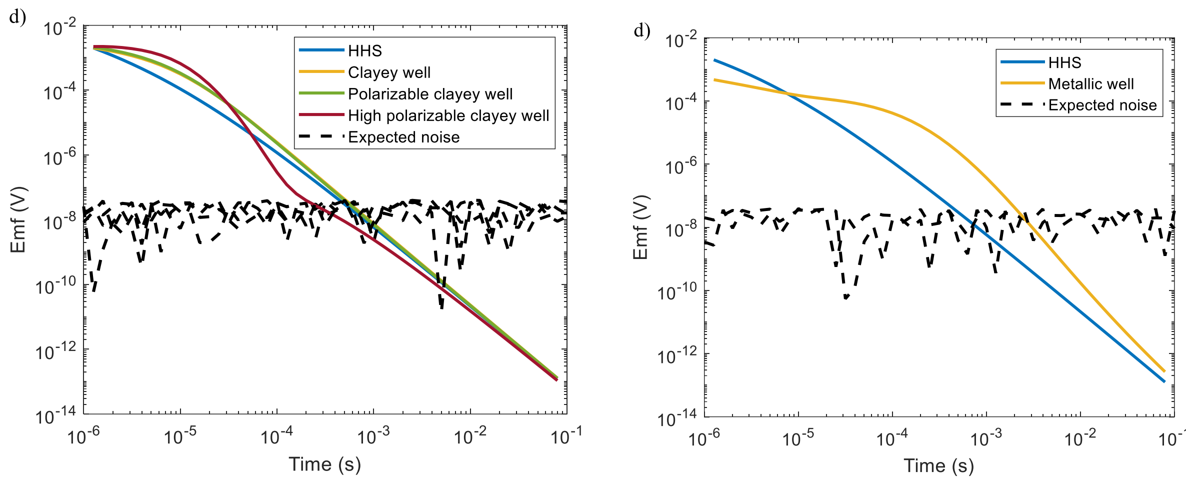


Figure 4.2. Drone TEM responses at 40 m flight heights for non—metallic, clayey (left) and metallic (right) wells. Homogeneous half space response (HHS) is shown for reference.

4.2.3 Ground Penetrating Radar (Priority 3, LLNL Lead)

<u>1D Electromagnetic modeling for expected well response:</u> In parallel with full 3D modeling described below (section 2.2.4) a low-cost, fast 2D surrogate was rapidly developed to mimic airborne GPR acquisition and study sensitivity of GPR to uncased wells (Figure 4.3). The model is based on finite difference time domain wave propagation technique. Various parametric simulations were conducted to determine the effect of soil conductivity (GPR suitability), buried well depth location, antenna offset location relative to well, drone flying height and presence and absence of casing to determine key experimental parameters for performing successful field testing as well as determine the feasibility of this approach. Key results are:

- Method can detect wells in low-conductivity soil (increasingly difficult with conductivity > 0.1 S/m),
- Observed weakening signal strength with increasing well depth (>1.5 m not detected) from attenuation,
- Method can detect wells not exactly perpendicular (or offset) to the GPR, up to a few meters,
- Sensor can be flown at heights up to 10 m from the ground surface for reliable detection.

To the extent that 2D simulations allow, these results are based on reasonable assumptions of ground and well material properties. The model includes 15 m of air layer on top of a 15 m deep subsurface with a 15 cm diameter uncased well, 3 m deep. The sub-surface is heterogeneous medium with dielectric permittivity mean 4 and standard deviation (σ =0.25) and conductivity mean 0.01 and standard deviation (σ =0.25) (Figure 4.3). A 300 MHz air-borne antenna pair is flown at 5 m above ground centered above the well. Demonstrating a clear scattering and reflection off the well. The observed radargram clearly shows the direct source coupling, ground reflection as well as the well head reflection.

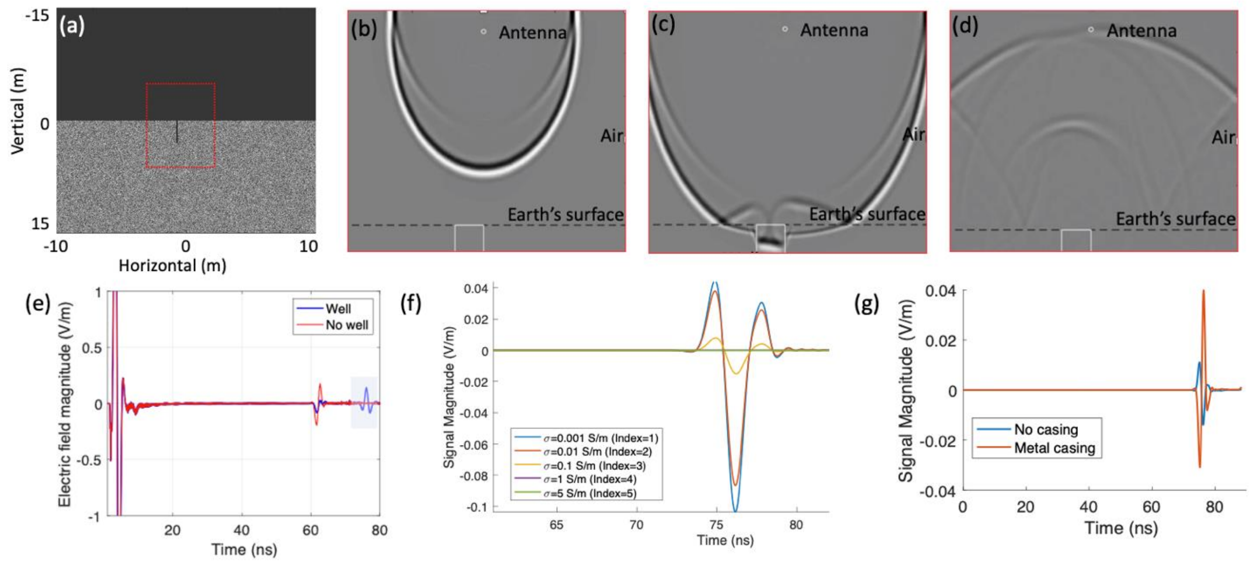


Figure 4.3: Surrogate 2-D numerical model to determine sensitivity of drone-GPR to well identification; (a) Dielectric map of the 2-D model, (b-d) electric field snapshots at 12 ns, 24 ns and 36 ns showing waves hitting the ground and scattering from the well and reflecting back to the antenna, (e) radargram can detect the presence of the well along with source contribution and ground reflection, (f) sensitivity of GPR signal to soil conductivity (GPR suitability index ranging from 1-5), (g) sensitivity of GPR signal to lack of casing.

Hardware development: A thorough market survey was conducted for state-of-the-art GPR systems based on several factors such as compactness, low package weight, sensitivity to non-metallic well casings and ease of integration with LLNL drone systems, resulting in procurement and assembly of the Sensors and Software Inc. pulseEKKO system. The GPR system has multiple resistive dipole antennas with an operating bandwidth will be between 250 MHz to 1000 MHz to optimize the spatial resolution required by target size and exploration depth. The hardware provides access to raw data and is integrated with GPS and odometers for accurate spatial positioning. Its fully bistatic design enables variable antenna offsets and orientations for advanced survey types for practical field operation. The GPR system was initially configured on a movable cart to take real-time measurements (Figure 4.4). The system was tested in Summer/Fall 2023 over concrete platform and successfully recorded reflections from the rebar within. Integration of the system with LLNL's drone platform **(WP3)** is planned for Winter 2023 and would lead to conducting the first-ever field demonstration of GPR technology for orphaned well detection.

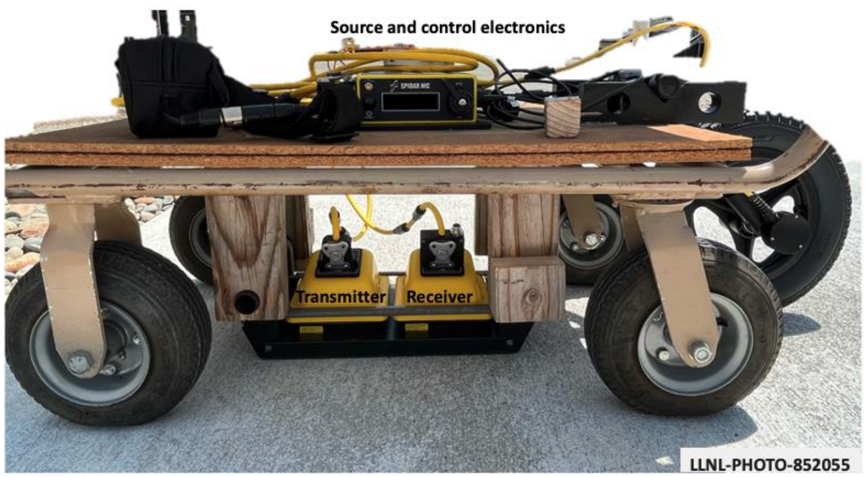


Figure 4.4: GPR system (250 MHz-1 GHz) integrated on a cart platform at LLNL scanning over a concrete platform (Frequency chosen strategically for deep penetration (1 meter) with < 1 cm resolution); (https://www.sensoft.ca/products/pulseekkopro/overview-pulseekko/)

4.2.4 Rapid, Scalable 3D Computational Modeling (Priorities 2-3 SNL Lead)

In year 1 of the project we modified the mixed vector/scalar potential code (Weiss, Computers and Geosciences, 2013) from its prior formulation as Cartesian finite volumes, to unstructured finite elements in anticipation of incorporating the HiFEM concept. And while initial testing of this potentials-based code yielded favorable comparisons with independent reference solutions, we were unable to successfully manipulate this formulation in a way that was HiFEM compatible. In response to this, we instead redirected toward a new formulation based directly on the electric fields (the "curl-curl" equations in the EM literature) and utilizing a novel set of "edge-based" (Whitney, or Nedelec) vector finite elements. While these edge-based elements came at a slightly higher cost than the nodal elements used in the potential formulation, we were able to use much of existing finite element library from the latter, thus accelerating development and testing time. Comparisons between the new broadband HiFEM EM code and reference solutions are favorable, as are internal consistency checks (Figure 4.5).

To ease use of the HiFEM ecosystem of multi-physics simulation into the workflow of end-users tasked with well identification and characterization, we note that HiFEM is inherently modular in its design (Figure 4.6). Mesh generation is done independently, outside of HiFEM, and thus users are free to choose their mesh generator of choice. In place is a translator (inp2cjw) that converts Abaqus-formatted mesh files to the HiFEM standard. Likewise, HiFEM outputs are VTK-formatted and thus readable by any number of data visualization packages. Cooperative agreements are already in place for HiFEM sharing among national laboratory analysts (for the case of notably difficult well identification scenarios), and **Sandia National Laboratories actively engages in software licensing to facilitate rapid market utilization and commercialization.**

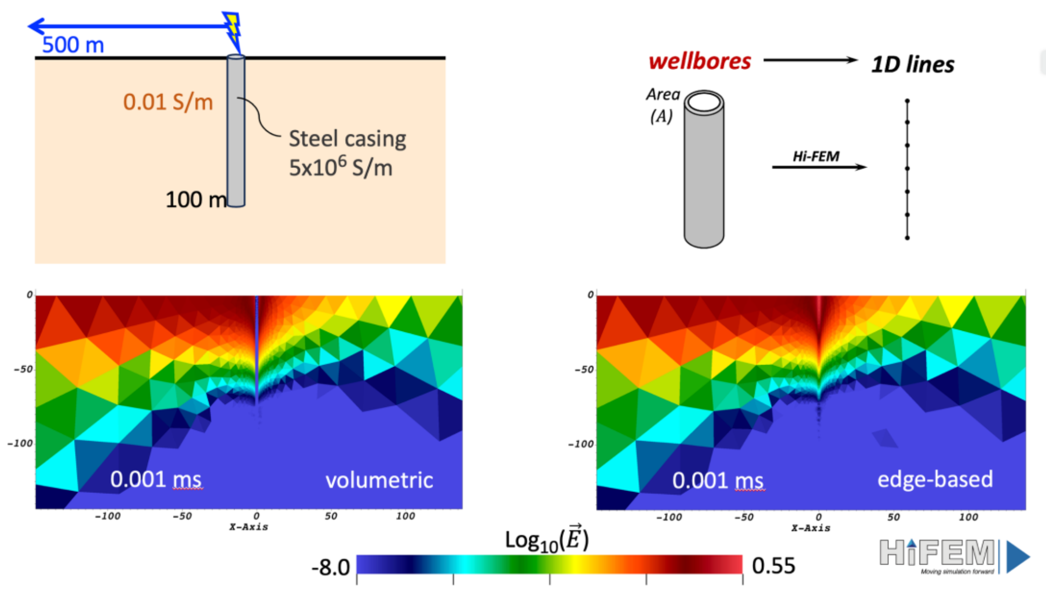


Figure 4.5. HiFEM validation exercise demonstrating the equivalence of volume-based (expensive) and edge-based (cheap) well representations for a top-casing stimulation scenario.

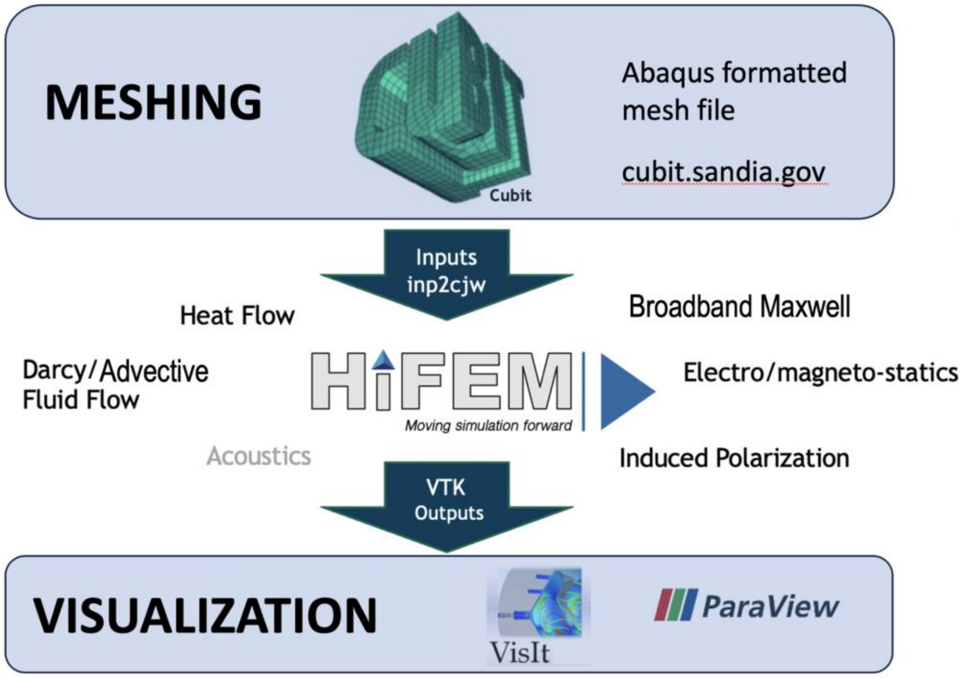


Figure 4.6. Modular design architecture of HiFEM multi-physics ecosystem whereby meshing and data visualization workflows communicate with HiFEM through standardized I/O progams.

4.3 Motivation and Background

4.3.1 Magnetometry

The presence of steel casing from well completions has widely been exploited for identification of abandoned wells through collection and analysis of magnetometry data. The large magnetic anomaly in these data is fairly diagnostic when the well is present, but absent when the well is uncased or cased by non-magnetic materials (e.g. wood). Furthermore, modern magnetometers of requisite sensitivity are cheap and small, thus accelerating their adoption by the commercial sector for airborne data collects, either by UAVs or manned platforms. Thus, there is a growing volume of magnetometry data both in-hand and soon to be collected, which requires defensible guidelines for its curation and analysis if the overwhelming number of undocumented wells are to be identified. The activities here focus on this immediate need for practical solutions to managing and collecting magnetometry data **(WP9)**, including its integration with alternate data sources and analysis frameworks **(WP3).**

4.3.2 Time- Domain and Induced Polarization Electromagnetics

A drone-based wellbore finding system using active sourced transient EM technology offers a few key advantages over magnetic tools, including (1) Easiness to use/interpret the data; (2) Sensitivity to non-ferrous materials; (3) More accurate location identification and (4) Less sensitivity to cultural noises. Developing such as a system provides not only an additional identification modality on top of airborne magnetic surveying, but also an opportunity to conduct-drone based surveys for the identification of non-metallic cased wells by identifying well pads/contaminant related signals using EM derivable signals, such as induced polarization (IP).

4.3.3 Ground Penetrating Radar

While magnetometry is the “go-to” method for identifying wells with metallic casing, it’s well known to be insensitive to the presence of uncased wells (wells whose casing were moved historically) and non-conductive casings such as wood. Back-of-the-envelope “Fermi estimates” from historical rig count data of well salvaged during World War II indicates that about 1600 – 15000 orphaned wells with casing pulled remain unidentified. To aid in their identification, we explore the feasibility of drone-based ground penetrating radar (GPR) systems, with support from both field trials and numerical simulation.

GPR employs radio waves (1-1000 MHz frequency range) to map structures and non-conductive or dielectric features buried in the ground [1]. While historically, GPR has been mostly used as a surface or near-surface based geophysical interrogation technique thus limiting survey speeds and scan areas, recent developments towards light-weight GPR hardware have led to drone-based GPR systems that can be flown up to a few meters high from the ground level. These include GPR-based drone platforms for detection of sub-surface buried threats in the soil and soil moisture mapping correlating accurate spatial GPR responses with soil moisture values and complex soil hydraulic properties [reference].

Taking inspiration from the recent advances, we are attempting to develop a cost effective and efficient drone-based GPR to identify uncased undocumented orphan wells. The depth of penetration and thus sensitivity to buried wells is highly dependent on the soil dielectric properties. Overlaying the GPR ground suitability map for the continental United States with the documented well location density map gives an indication where GPR will be most aligned with the well identification /characterization problem (Figure 4.7). The overlay results show the highest likelihood of success in northern Appalachia, with mixed results expected in mid-continent and south-central U.S.

1 + 2 (high) GPR suitability **1 + 2 + 3 (moderate) GPR suitability** **1+2+3+4 (low) GPR suitability**

Figure 4.7: Heat map of well density (red-high, green-low, white-none) filtered by GPR suitability index (1-most suitable, 5-least suitable). Most of the well locations lie in regions with moderate to high GPR suitability (middle panel).

4.3.4 Rapid, scalable 3D Computational Modeling

To support the data collects described in 2.2.1-3 described above, **WP2** focused its efforts around optimization of electromagnetic (EM) simulation tools. With few exceptions, the operational scenarios for well identification will be fully three-dimensional (3D) in their character, possessing irregularities in topography, variable geology and the presence of anthropogenic clutter outside of the well itself. This necessitates 3D modeling analysis (a solved problem), but more importantly, real-time or quasi-real-time model analysis (an unsolved problem). To meet this need, the SNL team pursued modification of its real-time electrostatic modeling tools [22,] for broadband (full-frequency) Maxwells Equations. **Success in this area impacts priorities 2-3, forms a basis for Earth/well characterization (WP4), and provides a viable tool for building ML training sets to inform multi-sensor data integration (WP3).**

To better understand the severity of the computational problem, a primary design feature of finite element models in which spatially-extensive, electrically-conductive features are present is the disproportionate number of computational resources needed for their representation in. a discretized finite element mesh. For example, discretization of steel well casing requires roughly 10M finite elements per km, a number which renders simulation of its electromagnetic response computationally expensive and typically requiring massively parallel-compute architectures for realistic field scenarios. To minimize this extreme computational burden, Weiss [21] proposed a hierarchical material properties distribution (Figure 4.8) whereby electrical properties of the finite element discretization can be attributed to not only the volume part of the finite element (as is commonly done), but also its facets and edges. Thus, wells, can be represented by a connected set of edges within the mesh, rather than many millions of tiny tetrahedral elements, thus reducing the computational resource requirement by several orders of magnitude. This not only allows for quasi-real-time model evaluation, but also minimizes the principal bottleneck to data inversion: the high cost of the forward problem. The flexibility of the finite-element discretization (HiFEM, for hierarchical finite elements) enables accurate representation of real-world oilfield scenarios based on aerial photography and as-built completion designs (Figure 4.9). Numerical examples such as these also inform novel experiment design concepts (and data collection strategies) whereby *in situ* infrastructure can be exploited to illuminate previously unidentified wells.

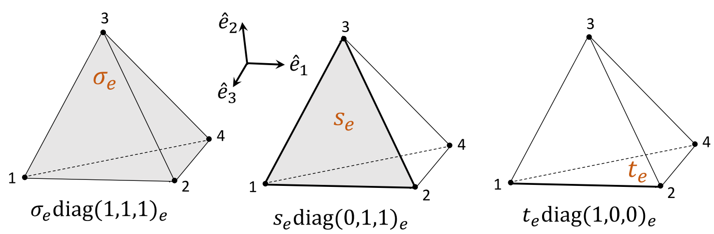


Figure 4.8. In the hierarchical materials properties framework (Weiss, 2017), electrical properties can be attributed to volumes (left), facets (middle) and edges (right) of the underlying discretization, thus minimizing the cost of representing thin features within a given simulation.

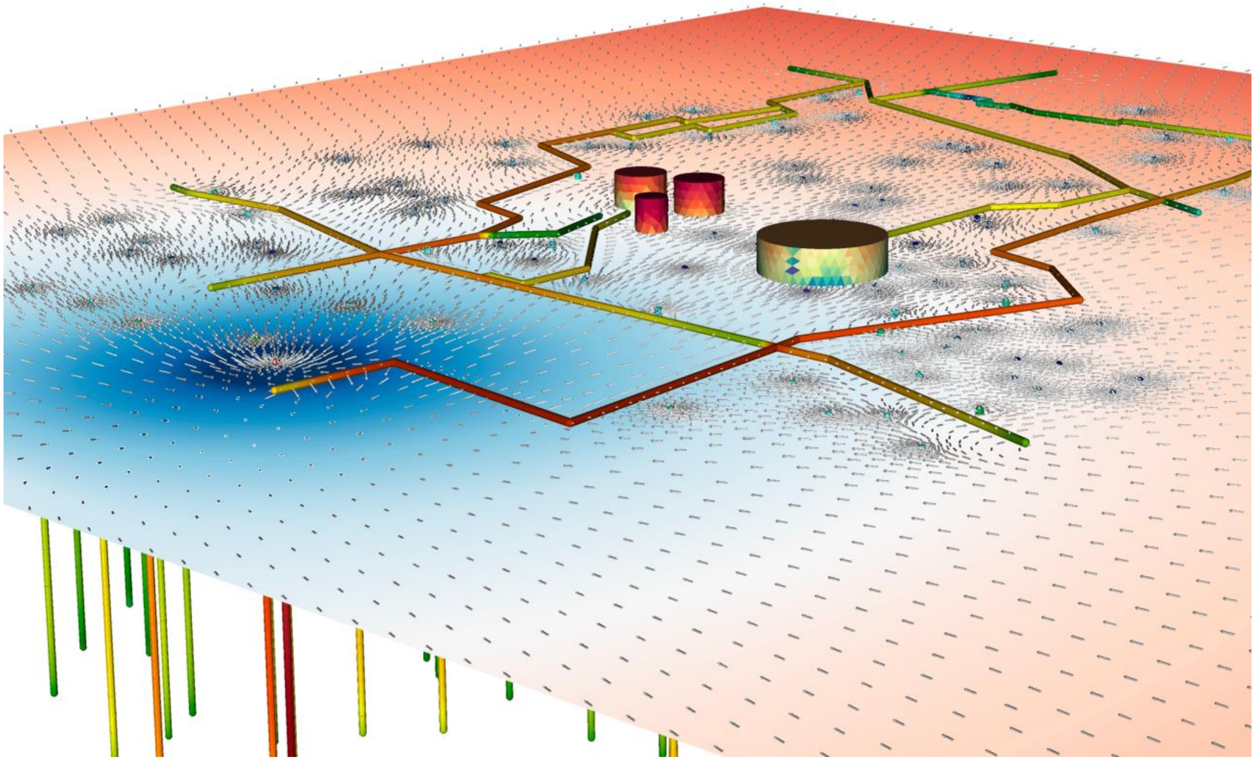


Figure 4.9. Numerical simulation of the DC response of a section within the Kern River oilfield [*] whereby top-casing electrification of a single well head (blue region, left) results in secondary electrification of nearby wells and top-side infrastructure such as the distribution pipelines and storage tanks shown here.

4.4 Data Summary

Drone-based IP data were collected by LBNL. Mag data were collected by NETL. GPR, IP, and EM modeling exercises were conducted at LLNL, LBNL and SNL, respectively.

4.5 Future Work

4.5.1 Magnetometry

- Test ML magnetic compensation algorithms at Stonewall Jackson Reservoir
- Evaluate new QuSpin total field and vector magnetometer at test area in Hillman State Park
- Expand database of field sites containing uncased wells.

4.5.2 Time- Domain and Induced Polarization Electromagnetics

- Complete numerical simulations of DroneTEM surveys with the aim to aid hardware design and optimization.
- Develop EM based IP numerical data simulations and inversion capabilities in support of the new system.
- Complete the prototyping and initial testing of the new DroneTEM system capable of higher flight height and acquire high quality transient EM data for wellbore identification (both metallic and non-metallic).

4.5.3 Ground Penetrating Radar

- **<u>Field measurements</u>**: The hardware is currently being integrated in an UAV platform to scan buried well regions, focusing on survey design and detectability under conditions of dense vegetation and irregular topography. These experiments will reveal practical sensitivity to wells without casing or with non-conductive, non-magnetic casing, and will be coordinated with **WP9.**
- **<u>Simulation Development</u>**: As programmatic needs require, the 2D surrogate numerical model will be extended towards 3-D to obtain more accurate GPR signals from complex sub-surface terrains. These models will be used to obtain tighter bounds and detection limits of wells without metal casing in sub-surface regions. **This will be benchmarked against the SNL HiFEM model.**
- **<u>Localization algorithms:</u>** A localization algorithm based on the electromagnetic time reversal focusing technique will be developed to back-propagate the simulated and measured signals and highlight or identify the location of the well. While conventional localization techniques rely on an antenna array, this algorithm would be able to localize the well locations based on a single transmitter-receiver pair dynamically moving in a drone platform.
- **<u>Machine learning:</u>** A machine learning technique will be developed that can incorporate the measured field radargrams and de-noise background clutter from multiple targets such as vegetation and targets that are not wells [4]. This algorithm will be capable of not only denoising clutter from the measured signals, but in the future will be employed for classification of wells. Finally, efforts will be conducted to scale and apply this algorithm towards other well identification techniques such as low frequency electromagnetics (SNL).
- **<u>Potential future work with additional funding:</u>** Leverage additive manufacturing to develop 3D printed light-weight microwave antennas with radiation efficiency comparable to more bulky current generation antennas, for ease of placement in a SAR or GPR identification system.

4.5.4 Rapid, scalable 3D Computational Modeling

- <u>Final validation and verification testing for model accuracy</u>. Compare HiFEM results with independent reference models to quantify the reliability of modeling results as a function of well/environment complexity, data acquisition design, and numerical discretization.
- <u>Coordination with LBNL and LLNL for data acquisition scenarios and parameters.</u> Determine which data streams will be available for subsequent modeling and interpretation. Document format standards and protocols for data exchange and curating.
- <u>Assert Copyright to enable code dissemination among end-users.</u> This is the first step in getting code into the hands of end users. With copyright asserted, terms of licensing can then be addressed.
- Coordination with **WP3** for ML learning set analysis. Because of the comparatively high computational efficiency of HiFEM, there exists the potential for building a ML learning (training) set from a collection of HiFEM model runs. This will require coordination with WP3 to document data standards/formats and data curation for their ML algorithms.

5. Work Package 3- Sensor Fusion and Data Integration with Machine Learning

5.1 Section Summary

Work Package 3 in the CATLOG program area is focused on sensor fusion and data integration with machine learning (ML). The 3 primary priorities for this year are:

1. Development of new integrated hardware platforms with multiple instruments
2. Development of new algorithms to identify UOWs from integrated pre-existing datasets (e.g. historical images, state databases etc.)
3. Development of new algorithms to identify UOWs from multiple remote sensing data sources

Significant progress has been made towards these priorities. LLNL is the primary team leading priority 1 in coordination with WP1 and WP9. LBNL and NETL are working on priority 2 focusing on identification of UOWs from USGS historical topographic maps using state-of-the art computer vision algorithms compared with current well inventories in state databases. LANL and SNL are working on priority 3 focused on data fusion and geospatial analysis of satellite, aerial photography, drone images from CATALOG field campaigns. As a cross-cutting work package, coordination with other work packages including WP1, WP4, WP5, WP7, WP8, and WPclassi9 has been an important aspect of WP3 activities.

5.2 Activity

5.2.1 Gap analysis, developing requirements for hardware, data integration and ML (Priorities 1, 2, 3)

5.2.2 Hardware integration (Priority 1)

5.2.3 Evaluate data integration approaches (Priority 2 and 3)

5.2.4 Selection, training, deployment, and benchmarking (ML and other) of algorithms (Priority 2 and 3)

Key Accomplishments for each activity are summarized below:

Activity 1: Identified data gaps in California state databases and working with the CALGEM state agency to obtain a full set of all well records in the state. Tools developed under the new work package 7 will be used to extract information from these well records for further analysis.

Activity 2: All first prototype sensors have been procured and many have been received. First multi-modal sensor system prototype is currently being integrated.

Activity 3: Released an integrated dataset consisting of private and public well data sets and select attributes (age, type, depth, etc.) on EDX (https://edx.netl.doe.gov/dataset/co2-locate). This activity has been spun off into a new work package (WP8).

Activity 4: Applying computer vision models to historical topographic maps and aerial photos to predict well location. Initial deep learning-based U-net model has potential to accurately identify well locations at various locations within the United States from USGS topographic maps.

Activity 4: Testing the utility of classical statistics and ML to predict well locations based on various geospatial and remote sending data types.

5.3 Motivation and Background

Multiple datasets need to be combined to create high-fidelity, probabilistic maps of undocumented well locations. These data can range from historical records and production records located in various state databases to data collected in the field by the other work packages that need to be synthesized and analyzed. Challenges associated with this task are to harmonize information across disparate data sources, and to develop algorithms that can pinpoint well locations and their characteristics based on the limited data that are available. Significant cost-savings can be achieved through sensor fusion onto single platforms.

All laboratories involved in the CATALOG project are engaged in the activities of Work Package 3. Responsibilities for each lab are shown in Table 5.1.

Activity #	LANL	LBNL	LLNL	NETL	SNL
3.1					
3.2					
3.3					
3.4					

Table 5.1: WP3 Activity Leads (Blue) and Activity Partners (Tan)

5.4 Data Summary

5.4.1 Gap Analysis, developing requirements (hardware, data integration and ML)

LBNL- Identifying data gaps and stakeholder engagement in California

Information about oil and gas wells from state agencies can be broadly separated into collections of original documents (e.g., well logs, field maps, etc.), and databases that contain information extracted from such documents. However, state databases often have missing information that are critical to identifying and characterizing wells such as well depths and spud dates. Given the focus of the LBNL team in California in Year 1, we examined the information present for all wells in the official well database issued by CALGEM. Through this analysis, we found that about 2/3 of the wells do not have a spud date in the database (Figure 5.1), which is information that would be useful to identify patterns in historical well drilling. We demonstrate the potential for using spud date information by using the currently available data and clustering wells according to their spud date. This reveals a clear temporal pattern of the evolution of drilling activity (Figure 5.1) that could be enhanced if more spud dates would be made available. The geographical distribution of such data gaps in spud dates is heterogeneous, with Southern California counties displaying more missing spud dates (see Figure 5.1). A check of a few records that were missing spud dates revealed that the information was present in the well records, indicating that this data gap could be overcome with additional effort to extract information from well records.

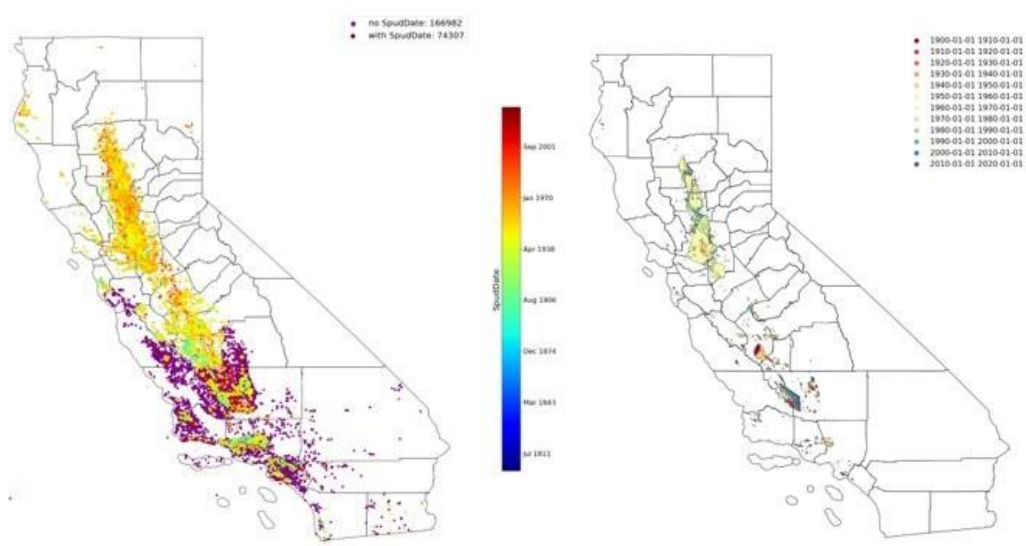


Figure 5.1: Maps showing wells organized by spud date. (a) About ⅔ of the wells in the CALGEM All wells database does not have information on spud dates. (b) Unsupervised spatial clustering of nearby wells spudded within the same decade indicate potential patterns in well drilling.

The data about oil and gas wells collected by state agencies like CALGEM constitutes a valuable resource of publicly available information for the discovery of undocumented wells. However not all the information contained in the original documents are properly extracted and made available to the public.

We contacted CALGEM requesting their full set of digitized well records in the State of California, prioritizing previous orphaned wells. These records are in the process of being transferred to LBNL. Having access to the original documents will allow us to conduct further work on information extraction using Optical Character Recognition, Natural Language Processing and Large Language Models in coordination with WP7.

Another opportunity for filling data gaps is by comparing state databases with third-party datasets. Here we compared the geographical distribution of wells compiled by the USGS using IHS Markit data, where well counts are aggregated using a 2-mile side mesh, with the well locations from the CALGEM database. Figure 5.2 shows a comparison of the two distributions as a count difference, both as a geographical map and as a histogram. The map shows again some heterogeneity in the differences, consistent with the geographical domain of the missing spud dates. The histogram shows that even though most of the areas contain a comparable number of wells (note the logarithmic scale on the y axis), there are locations where this discrepancy is significant. Such areas may represent a starting point for addressing data gap filling.

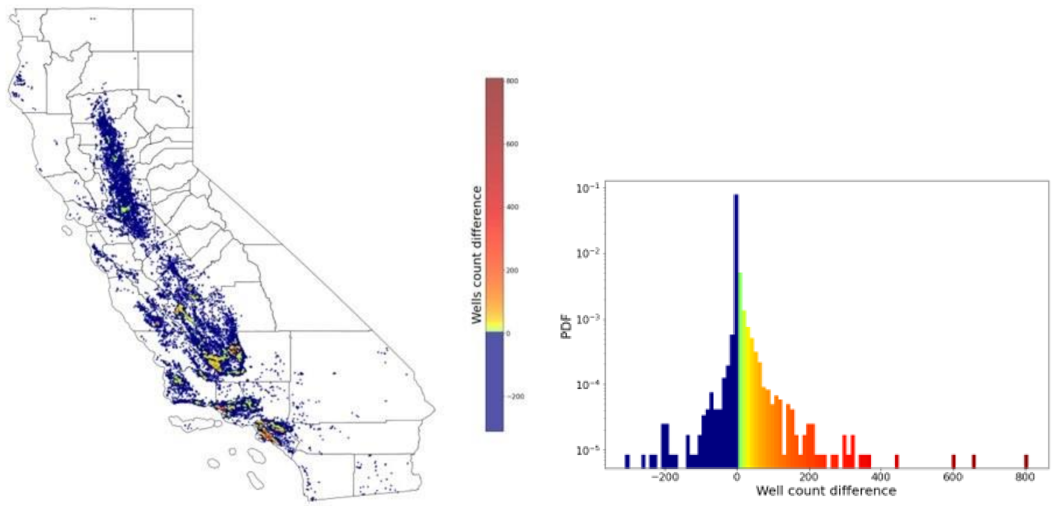


Figure 5.2: Well count difference per 2-square mile between the IHS aggregated dataset and CalGEM.

These preliminary findings justify the effort, in collaboration with new work packages 7 & 8, to improve information extraction from original documents in order to fill these gaps in an effort to construct a national database of known and undocumented wells. The workflows we have developed for this analysis is not limited to California and are extendable to the entire US.

Identifying data gaps and stakeholder engagement in Oklahoma (SNL)

Contacted Oklahoma Geological Survey

5.4.2 Hardware integration

LLNL performed technology assessment of sensors to better understand system requirements and the value proposition of different combinations of sensors. The table below shows the different sensors and the assessment areas looked at for each sensor.

	Stereo Camera	HSI Camera	Magnetometer	LIDAR	GPR	Gas Detector
Detection Score	0.1875	0.28125	0.390625	0.046875	0.71875	0.0125
Percentage of wells	1	1	0.5	0.5	1	0.1
Directivity	0.1875	0.28125	0.78125	0.09375	0.71875	0.125
Primary Quality	0	0	1	0	1	0
Secondary Quality	0.5	0.75	0.5	0.25	0.25	0
Tertiary Quality	0.5	0.75	0.25	0.25	0.25	1
Standoff (ft)	100	300	30	300	3	300
Scan speed (ft/s)	30	15	10	15	10	30
Processing Requirement (TFLOPS)	2	30	1	1	5	0
Size (in^3)	9	54	864	64	6480	12
Weight (lbs)	0.37	1.5	4	2	10	1
Power (Watts)	2	20	5	10	20	0.5
Cost ($)	500	30000	20000	5000	30000	1000
Total Utility Value	3829.69927	67.11306	1354.149	3290.216	425.8524	62.98273

Table 5.2 Value Proposition Matrix for Sensors

- LLNL has procured and received:
  - A high-standoff methane detector (Laser Falcon)
  - Ground Penetrating Radar (GPR) with multiple frequency options (PulseEKKO, 250 MHz, 500 MHz, 1000 MHz)
  - Stereo camera
  - Nvidia based processing system.
  - HESAI XT32 LiDAR (Topodrone LiDAR 200+)
- Aeris closed-cell methane detector (selected in coordination with WP1) have been procured but not received.
  - Multirotor sUAS parts have been procured and integrated; high-standoff methane detector and magnetometer have been integrated onto the UAS.

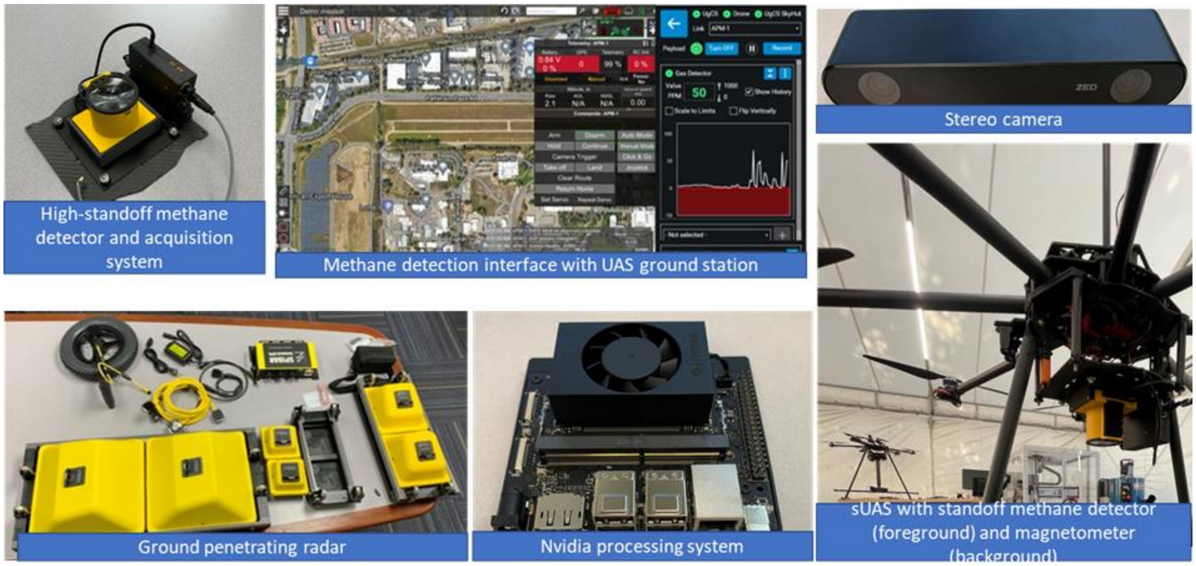


Figure 5.3 Procurement and Initial Validation for Sensors Selected

Initial data collections have been conducted with the GPR and Lidar. Below is an example of data collected with the Lidar mounted on a UAS, the Lidar was flown at an altitude of 250ft over an area with ground water sampling wells, the wells are constructed so 2 feet of casing are above ground level. LLNL is currently working on developing a method to improve the GPS location of the Lidar sensor to improve stitching together multiple passes of the lidar and reducing artifacts in the images such as an object showing up multiple times.

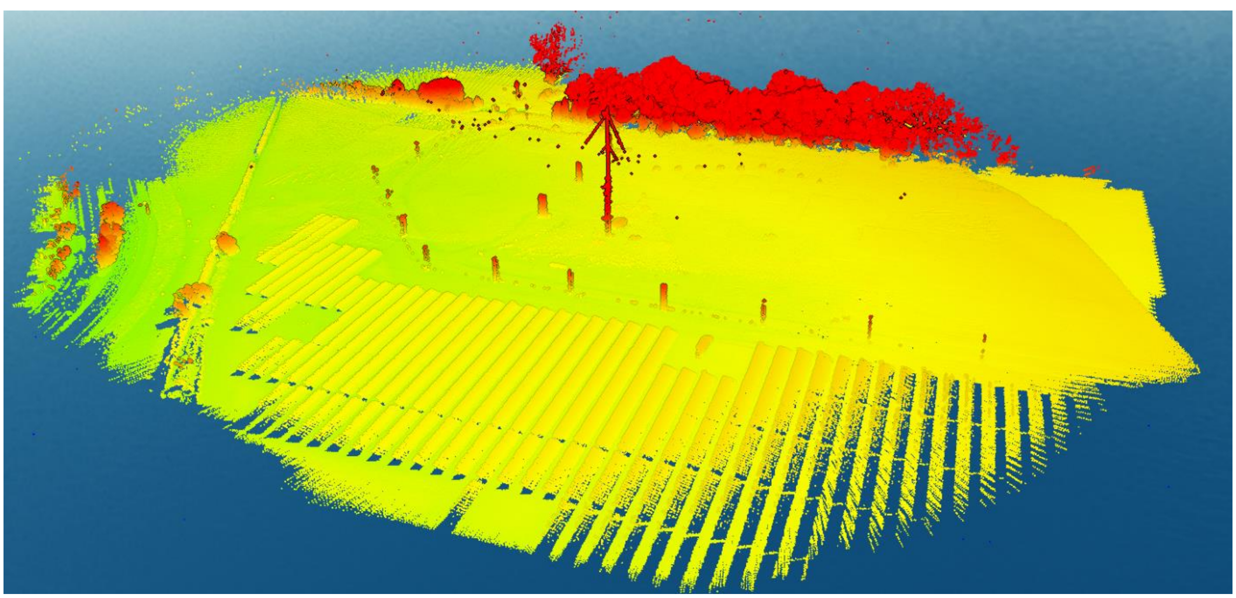


Figure 5.4: Image constructed by Lidar scans, coloring done by elevation, green lower in elevation and red higher in elevation.

5.4.3 Evaluate data integration approaches

Creating integrated data products for analysis and ML to Identify Wells

SNL

SNL spatial analysis efforts are focused in Oklahoma. We compiled spatiotemporal data that may contribute in identifying areas with a high- or low-likelihood of O&G wells being present. This began by combing the Oklahoma Geological Survey and Oklahoma Corporation Commission's data repositories as well as collecting subsets of remote sensing data with nationwide coverage (LandSat, National Agriculture Imagery Program, National Land Cover Database).

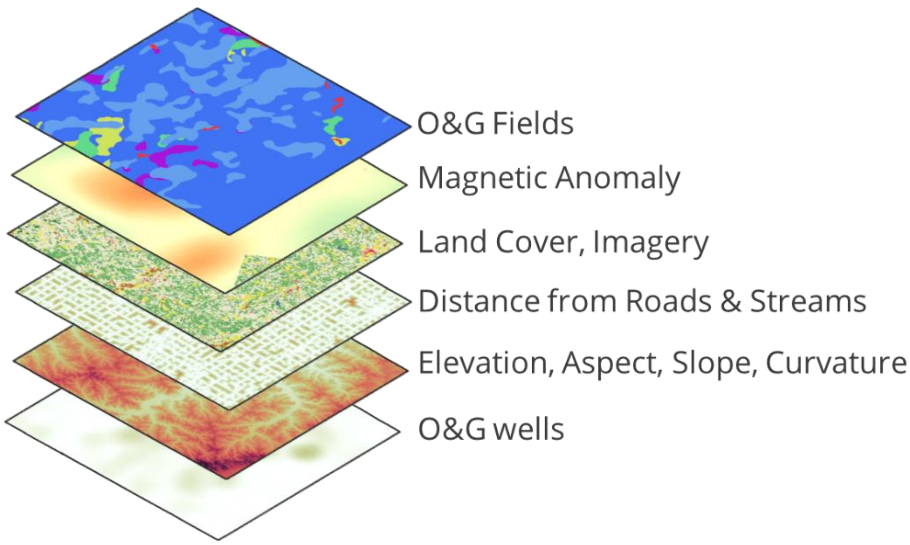


Figure 5.5: A range of environmental geospatial data are considered in creating a GIS model describing the likelihood of O&G wells being present.

Initial spatial analysis of the 324,250 O&G wells on record in Oklahoma (14,785 documented orphan wells) show wells to be highly spatially clustered (Figure 5.7). No clear temporal clustering in documented wells, orphaned or otherwise, is observed. At this stage we focus our modeling efforts on Lincoln County with the intention of upscaling later. Lincoln County is especially dense in active and documented orphan wells (Figure 5.6).

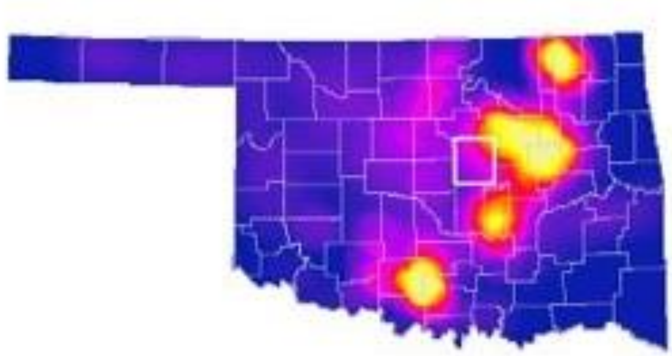


Figure 5.6: Heat map of O&G wells on record on Oklahoma (purple, low density; yellow, high density). White polygon delineates Lincoln County.

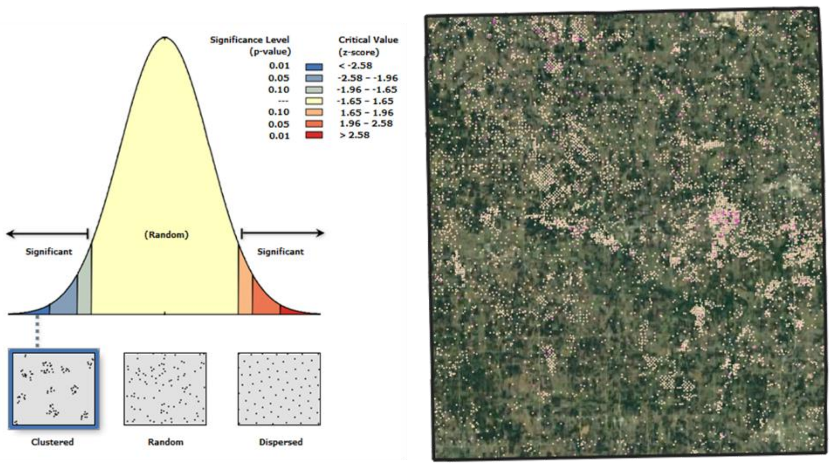


Figure 5.7: O&G wells on Lincoln County are highly spatially clustered (statistical distribution curve). No clear temporal clustering is observed. Peach dots represent 1 well in a 30m x 30m area; pink dots represent 2 or more wells in a 30m x 30m area. (Well Data Source: Risk Based Data Management System for the Oklahoma Corporation Commission.)

SNL is adopting two GIS approaches: multivariable regression (linear) and supervised deep learning (nonlinear). Both approaches are attempting to build a model using the same inputs and spatial resolution. Our independent variables are airborne magnetic anomaly data, Euclidean distances from roads, streams, and fault lines, and LiDAR data products (elevation, slope, aspect, curvature).

Multivariate regression started with an Ordinary Least-Squares (OLS) Regression where the dependent variable is well density per 30m-by-30m grid cell. Only the magnetic anomaly data has a weak linear relationship with well density. Considering the fairly coarse resolution of aeromagnetic data, compared to UAV and ground surveys, this can be a function of clusters of wells and a secondary association with other types of infrastructure or geologic anomalies overlapping with well clusters. (Of course, geological features that create magnetic anomalies may have a real correlation with favorable well production, but the focus here is on whether the well casing produces an anomaly that can be utilized to detect undocumented wells.) Figure 5.8 shows the results of OLS and demonstrates that environmental factors alone do not explain the spatial density of O&G wells. OLS residuals, though normally distributed (not shown), are highly spatially correlated. As a follow up check, we will perform a Geographically Weighted Regression (GWR), which essentially produces multiple OLS models across are geographical area, to determine if these environmental factors do have a linear relationship with well density in some areas and not in others. In any case, modeling certainly requires human geography inputs as well.

Figure 5.8: OLS estimated well density: Yellow, overestimate, white, good estimate, brown, underestimate.

To address the nonlinearity between our independent and dependent variables, we have begun a Random Forest (RF) Model. RF Regression is a supervised learning approach that employs an ensemble of learning methods and decision tree framework. The training sets contains true positives defined as documented O&G wells from the state's database and true negatives defined in the database as well records where the well was not drilled or there is no evidence of drilling. This introduces a data imbalance that is still being worked out (9,181 true positives vs 757 true negatives). Initial work has quantified the importance of our input variables (Figure 5.9). Again, magnetic anomaly has the greatest impact on the model.

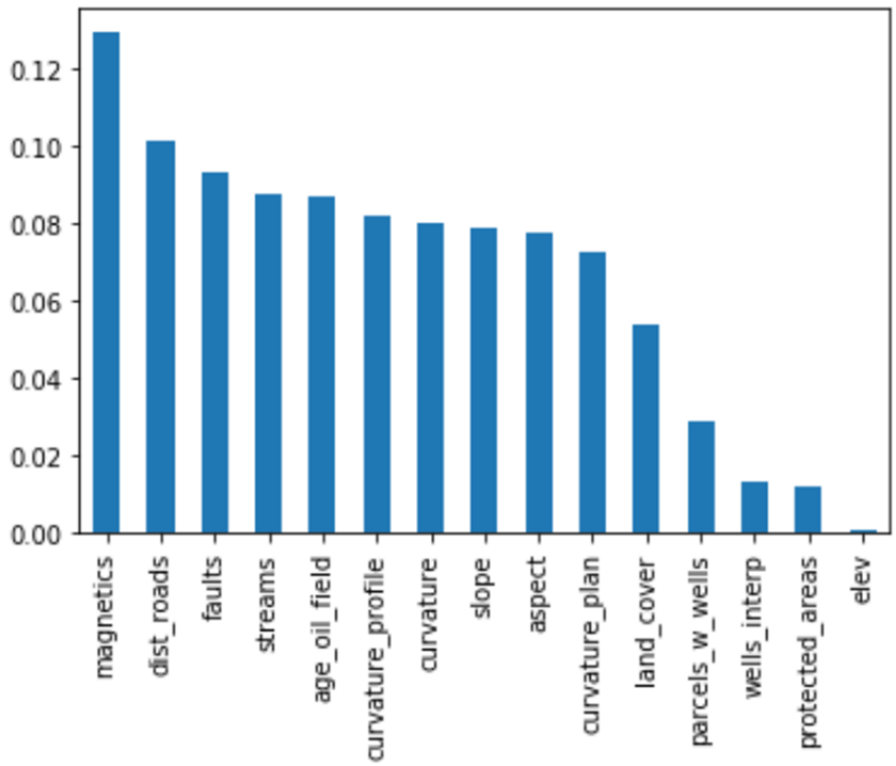


Figure 5.9: Bar chart of variable importance in Random Forest Regression. Importance (values relative to this specific model) is a measure of how much outputs change when a variable is omitted from the model. Magnetics is the aeromagnetic survey data.

Moving forward it is important that we incorporate human geography in our model as human behavior is clearly a major driver in the occurrence of a well. Specifically, conversations with stakeholders in OK pointed out that early (pre-geologic mapping) drilling was driven by economic opportunity (i.e., oil seeps or knowledge of productive wells nearby.) We identified census records as a source of information. Occupation entries, organized by town or neighborhood, can be categorized as O&G industry related or not and mapped through space and time. We can then determine the spatial relationship between these census entries and the well database. Currently, census records are searchable by place and names, which makes the data prohibitive to work with. Data scraping efforts are necessary to extract this potential dataset from scanned pdfs or images.

This work was presented at the 2023 International Meeting of Applied Geoscience and Energy.

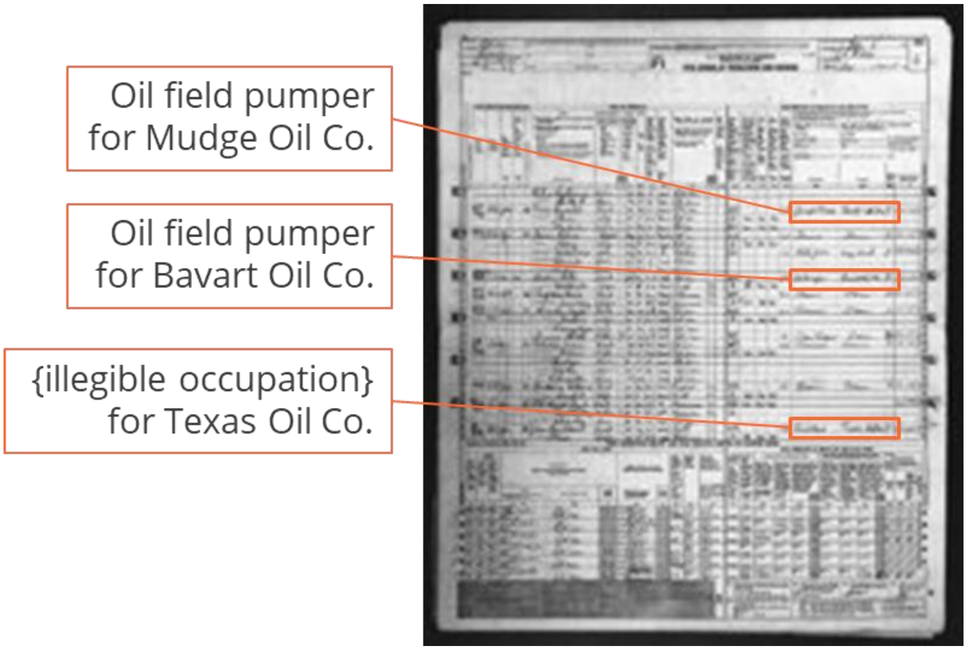


Figure 5.10: Example of census records with O&G industry-related occupations.

<u>SNL & LANL-Building a Satellite Image Database and Image Classification</u>

O&G wells are their associated infrastructure can often be recognized in satellite images or aerial photos and image classification can automate this otherwise labor-intensive work. LANL built a dataset of 120,000 cropped images from the National Agricultural Imagery Program (NAIP). Half of the images contain an O&G well using the database created by Boutot and others (2022); the other half do not contain a well. Associated with each image is a reference to the well record and the pixels within the cropped image where the well is located (Figure 5.11). This image set serves as the training set for the image classification algorithms used by SNL which produces a binary classification system (well or no well) as well as a pixel list of where the well is expected to be. To date, SNL's image classification has been focused on four counties in Oklahoma where tree canopy reduces confidence (Figure 5.11). The image dataset represents a variety of terrains so an additional piece of information will be a measure of ML suitability across areas.

The draft manuscript for this effort, specifically the creation of an image dataset, is complete. As a follow-on (phase 2, perhaps) to this work, we will build a database of historical aerial photos with the same objectives.

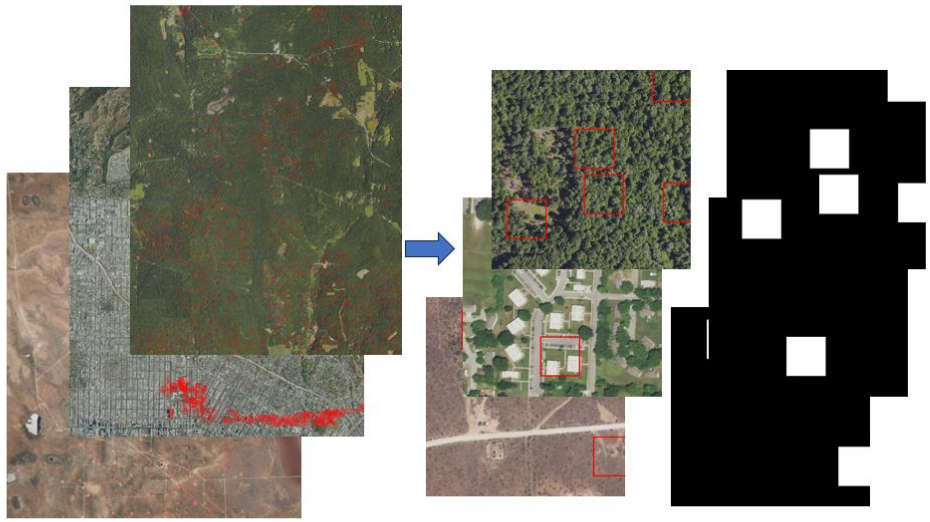


Figure 5.11: Examples of NAIP images that correspond with O&G wells from Boutot et al., (2022) database.

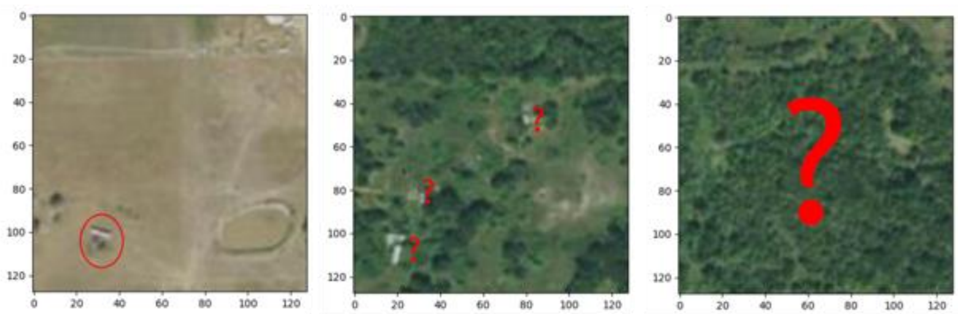
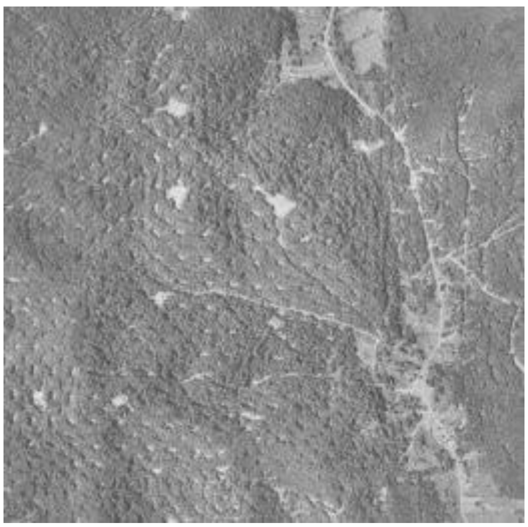


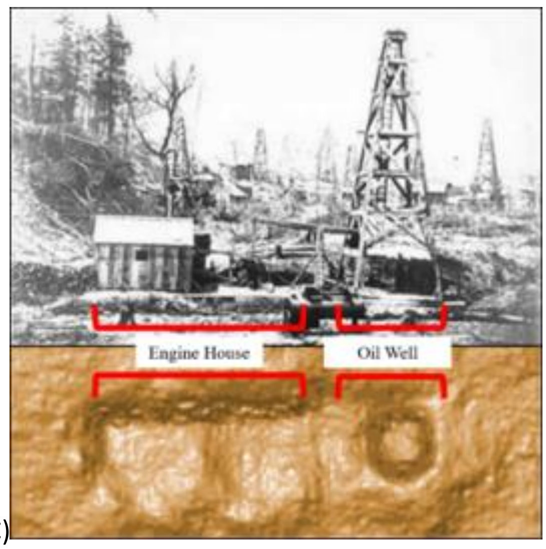
Figure 5.12: Image classification is highly dependent on the amount of tree canopy in imagery. Dense tree cover does not have a clear well signature.

Historical map digitization and compilation of other data sources (NETL)

Accomplished: Explored availability of, and began cataloging, historic maps with oil and gas well location information in Pennsylvania.

A)


B)


C)


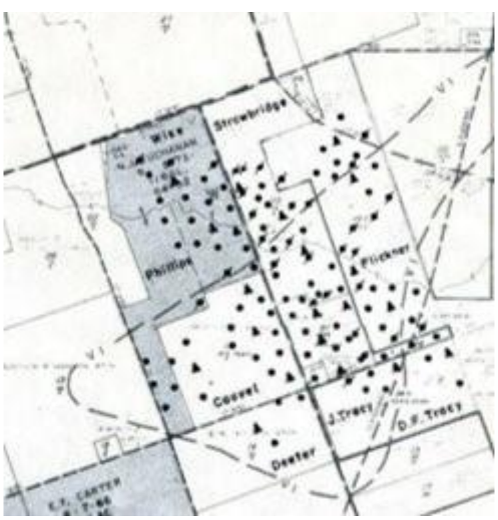
D)


Figure 5.13: Examples of historic documents and data sources that can portray or be used to confirm undocumented well locations, including A) USGS quadrangle maps, B) aerial photography, C) LIDAR surveys, and D) farm line maps.

Well database activity (NETL)

To aid in the identification of UOWs, the NETL team began working to expand upon previous projects and research activities to look at methods of integrating authoritative public data for wells from across the U.S. The NETL team started to build off work from the CO2-Locate database (Romeo et al. 2023), which has developed a baseline method and database schema for integrating geospatial well data from multiple, authoritative sources. Working with this baseline dataset, the team worked to integrate additional public well data resources from state and tribal regulatory authorities, as well as review and quality check the incoming data for normalizing attributes from across the different data sources (e.g., field names, units of measure, etc.). The project team also began to review database attributes in relation to the terms and standard definitions being developed by WP5. With the focus on UOWs, the team also worked to start integrating the USGS' Documented Unplugged Orphaned Oil and Gas Well Dataset (https://www.usgs.gov/data/united-states-documented-unplugged-orphaned-oil-and-gas-well-dataset) as well as the few databases from state regulators that have identified or flagged confirmed orphaned wells, such as the New York Department of Environmental Conservation's confirmed orphaned wells dataset. The project team also started looking at other available sources of data from the Appalachian Basin that could be used to verify or confirm undocumented wells, including the USGS quadrangle maps, aerial photographs, LIDAR, and farm line maps from the region as potential sources for future activities (Figure 5.13).

Currently, the WP team has integrated more than 50 publicly available data resources, providing information and locations on more than 4 million well records across the U.S. The team also processed national well data from proprietary databases, like Enverus and IHS, to provide derivative, spatial summaries for more than 4.6 million wells across the U.S., providing summaries related to wellbore spud dates, depth, status, etc. As of now, a subset of these data is available for public use and download through version 1 of CO2-Locate (Romeo et al., 2023; https://edx.netl.doe.gov/dataset/co2-locate). Version 1 database provides more than 2.1 million well records from public data sources, from Alaska, Alabama, Florida, New York, Pennsylvania, South Dakota, Utah, Virginia, and West Virginia, plus some national data from the Homeland Infrastructure Foundation-Level Data (HIFLD).

This goal and value of this database is to provide a summary of all available, authoritative data resources for well data available across the U.S. to help provide team members and the general public with insights into the similarities and differences in the data from each regulatory agency, as well as what additional information is provided by commercial well data sources. Additionally, this integrated resource aims to keep track of all sources of data, which has allowed for analysis to better understand the breakdown of well information provided from each data resource (Figure 5.14), including any potential differences in locations reported, as well as differences in the wellbore attributes (e.g., age, depth, operator name, status, etc.).

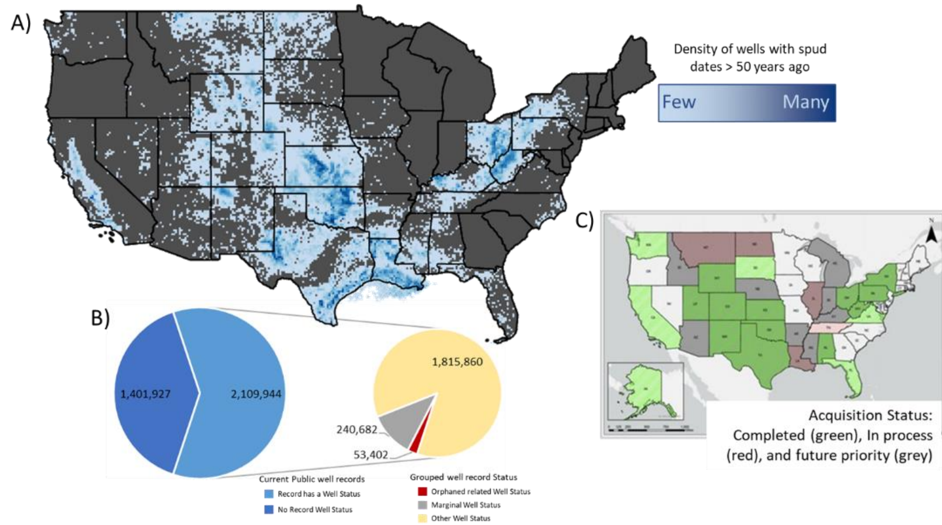


Figure 5.14: A) Density of wells with spud dates greater than 50 years ago as provided by the Co2-Lcoate, version 1 database. B) Breakdown of wellbore records based off status information provided, and if provided how those might breakdown and related to orphaned, marginal, or other types of wells. C) Data acquisition progress for national well database to spotlight states where public wellbore data has been collected and completed, is in progress, or has been identify for future needs.

5.4.4 Selection, train, deploy, and benchmark (ML and other) of algorithms

Well location probability assessment using historical data

We are developing machine learning based approaches for identifying UOWs and IOWs by extracting information on well locations from historical maps. In particular, topographical maps from the United States Geological Survey (USGS) published from 1884 to 2006 are available from the Historical Topographic Maps Collection (HTMC), a digital archive of approximately 190,000 georeferenced maps. These maps 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. These maps were chosen for our study due to their unique depiction of historic oil and gas locations, represented by circular hollow black map symbols. With a resolution of 1:24000, they ensure detailed representation. Importantly, these specific well symbols are predominantly found within the 1950s and 1960s subset of the HTMC, differentiating them from other chronological collections. 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 natural discolorations of the original maps after years of use.

We are using complementary computer vision models to identify the wells across different geographic locations and O&G development histories.

The LBNL team is focused on identifying possible locations of wells in California and Oklahoma that represent different types of regions (e.g., currently producing oil fields in Central Valley vs. wells in urban locations such as Los Angeles). We experimented with various algorithms to identify well symbols in historical maps including traditional computer vision methods like edge detection and template matching to more recent neural networks for semantic segmentation (Ciulla et al., in prep.). Our final approach uses the deep learning algorithm U-Net that has a precision of 0.98 on our validation dataset (O'Malley et al., in prep.). Once the wells have been identified, their location is compared with the ones present in the official databases 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 (Figure 5.15), and are marked as potential candidates for field investigators to verify the viability of using the topographical maps to identify UOWs. Initial investigations by the LBNL field team in the Central Valley and the Sacramento delta showed the potential of using this approach to identify UOWs, with further fieldwork planned through work packages 1, 4 & 9 to confirm the presence of wells identified by our algorithms.

Figure 5.15: a) Inset of historical topographic maps in Lincoln County, OK with detected wells highlighted by blue circles, documented wells in purple and UOWs in red. Blue numbers are unique ids, while burgundy values represent the distance in meters of a UOW from the closest documented well. b) The same area in a) shown from Google Maps. UOWs and documented wells are highlighted by the same symbols as in (a). In (c) a zoomed version of (b) allows one to visually distinguish the pump jacks.

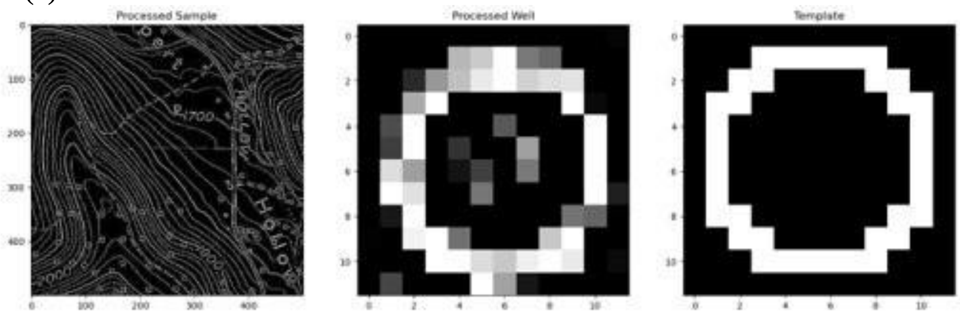
The NETL's team initial focus was on two key areas: Zachariah, KY and Olean, NY. Research conducted by NETL's field monitoring team has indicated that in these regions, certain well sites, despite field verification, remain absent from state-managed oil and gas databases. These regions were selected primarily because of the reliable, field-tested data provided by NETL's field monitoring team.

Considering the labor-intensive process of digitizing these points and the extensive number of HTMC maps relevant to key oil and gas regions in the U.S., the development of an automated machine learning model for feature identification is of paramount importance. To create the training data for our machine learning work, we digitized the quad maps and then extract the buffered digitized locations of the wells. The GeoTIFF format is an HTMC map format that can seamlessly incorporate into GIS applications. Due to blurriness and distortion, it was difficult to clearly distinguish symbols on some maps in the GeoTIFF format. In these instances, we referred to the map's GeoPDF version (HTMC's default format) to help discern the map symbols. Using Python, GeoPDFs were converted to GeoTIFFs, incorporating an upscaling step to preserve the higher resolution and clarity inherent in GeoPDFs.

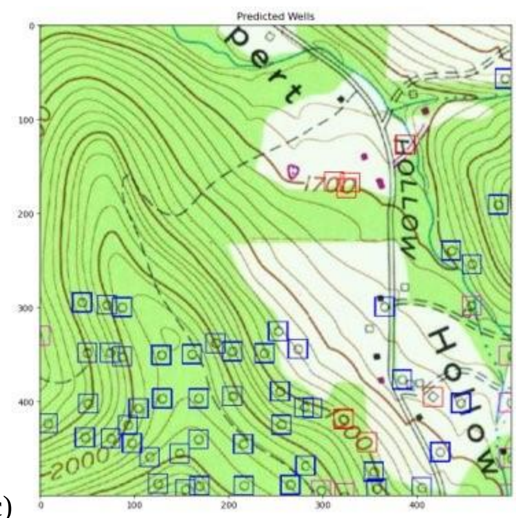
Initially, we explored an unsupervised template matching technique to detect the circular well symbols on the HTMC maps. The effectiveness of this approach was evaluated using the True Positive Rate (TPR) as a metric. However, this technique had its limitations, such as confusing zeros with well symbols and failing to capture wells located near the map boundaries.

$$TPR = \frac{TP}{TP + FN} \quad Precision = \frac{TP}{TP + FP}$$

(a)

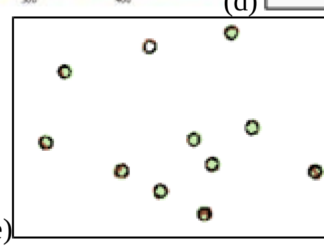


(b)



(c)

(d)



(e)

(f)

Current Template Matching Metrics	
True Positive (TP)	2830
False Positive (FP)	1166
False Negative (FN)	34

TPR	0.988
Precision	0.701

Figure 5.16: a) Formulas to calculate the True Positive Rate and Precision. b) Representative image of a circular map well compared to a template circle. c) Inset of HTMC map with detected wells highlighted by blue square and the false positives highlighted by red squares. d) ArcGIS model to create labelled training sets by creating a buffer around the map symbol and extracting the features. e) Labelled circular map symbols used to train and test the model. f) Template matching model metrics.

Together the LBNL and NETL teams investigated various approaches for well identification from the topographical maps and have converged on the U-Net model being a promising method for accurately identifying UOWs using the desired map symbols given its enhanced segmentation proficiency. The U-The algorithms developed by the two teams differ in training data/approach and implementation, which collectively indicate high precision in very distinct geographic regions. Future work under this work package will involve comparing these two approaches and sharing datasets to increase the confidence in well probability assessments using the topographic maps. The ultimate goal is to develop a scalable approach that can be applied across the continental United States.

LANL- Identifying area with high density of Orphan Wells (OW) using historical data and Geographic Information Systems (GIS) analysis

Project objectives: Identify the localized areas in a state with high density of Undocumented Orphan Wells (UOWs) so that we can prioritize the areas for future survey and advanced analysis using ML model, drone, magnetic survey, remote sensing etc.

In order to understand the intensities of the problems associated with UOWs, IOWs and DOWs and to select areas for future surveys and advanced machine learning and data analysis, we need to know how many UOWs we have across the US. Currently, there is no reliable estimate for how many orphan wells we have across the entire United States (US). Therefore, we have developed a workflow to estimate the number of UOWs, IOWs and DOWs by combining the available data. It comprises historical well completion, data digital records from corresponding state regulatory agencies and statistical correlation between the production and well completion where the number of well completion data is missing or not reliable. Using the workflow, the estimated minimum numbers (lower bounds) of UOWs for Pennsylvania and Oklahoma are 340,827 and 309,462 . This workflow can be applied to other producing states to estimate UOWs count on that state and we can get estimated total inventories across entire US.

The spatial distribution to the county is determined from the historical distribution of the known location of the wells (IOWs) in that region, and estimated UOWwell density is calculated. Using this method, we have found counties with a high density of UOWs using GIS in Pennsylvania and Oklahoma (Figure 5.17), which are very relevant to the development of boomtowns in that region. In Osage, Creek, and Okfuskee County has more than 25,000 UOWs in each county. And in Pennsylvania McKean County has more than 33,000 OWs and Jefferson, Armstrong, Venango, Clarion, Washington and Indiana County has 20,000-28,000 OWs in each county. The large number of UOWs in a county is associated with major fields that recorded huge production levels in the past and with secondary recovery techniques where waterflooding played a significant role in enhanced recovery. This distribution is immensely instrumental in identifying high-risk areas associated with many UOWs to prioritize those areas for future characterization.

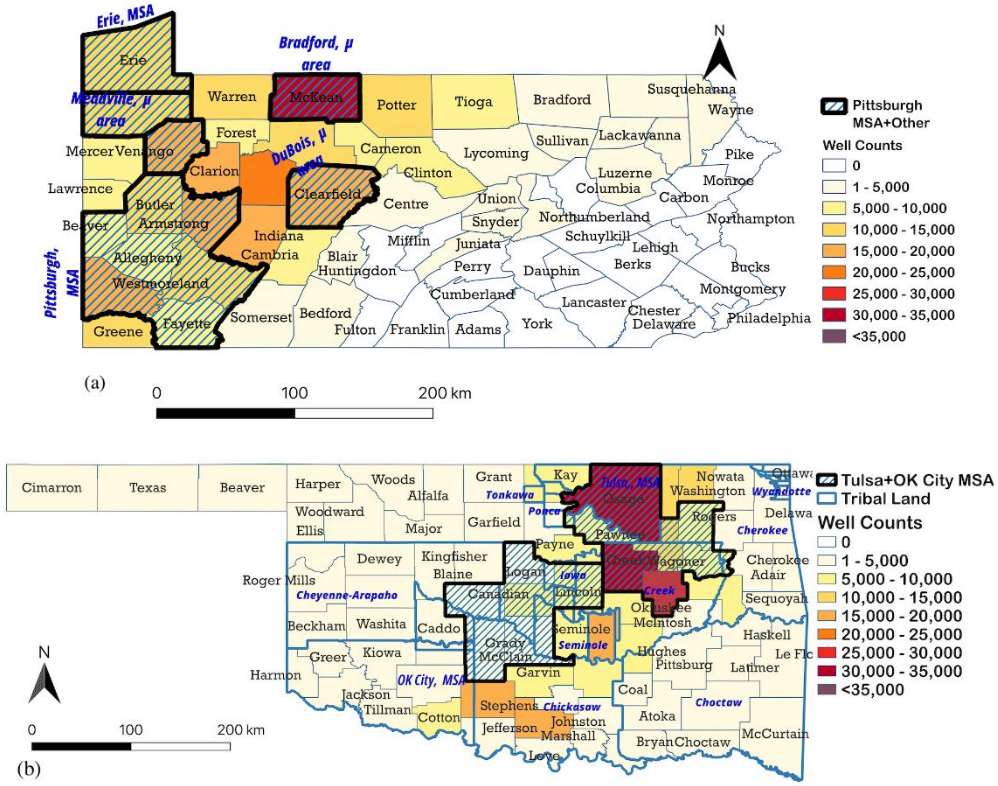


Figure 5.17: Estimated number of UOWs per county, including ER wells and potential underreporting in the past (a) in Pennsylvania during the entire state history 1859-2023 , and (b) in Oklahoma during the entire state history 1891-2023.

LANL-Well location probability assessment using multiple remote sensing and geospatial datasets

Transformer models have shown impressive results in fusing data from different sources (e.g., text and images). Our approach consists in simulating a drone flying with a magnetometer and a methane sensor taking measurements every couple of meters. These data are fed into the transformer model to reconstruct the well probability map as shown below. Our approach suggests that having two data sources (compared to just a methane sensor) increases the accuracy of the model by a wide margin, even when the additional data is noisy. A paper describing our machine learning approach has been accepted for *Nature Machine Intelligence*.

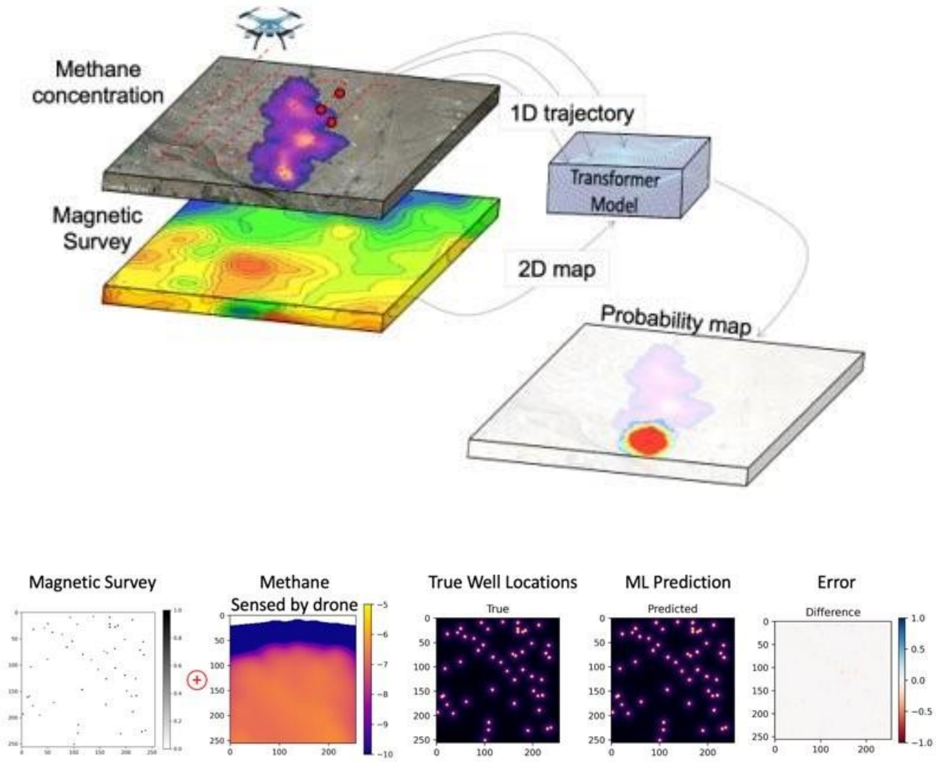


Figure 5.18: (Top) Diagram of the Process: Different types of signals (like methane and magnetic) are combined and fed into a transformer model. This model is trained to understand how these signals can show where wells might be in a certain area. (Bottom) Results: When we use both types of signals, the model predicts all the wells in our dataset with a high degree of accuracy.

To identify the location of wells using sensor data fusion, LANL developed an initial dataset and ML model with synthetic but realistic data for methane emission and magnetometer data based on information passed on from NETL at Hillman State Park. We are currently increasing the complexity of the datasets towards which we are applying our machine learning approach.

In a related effort, we have begun collecting a large dataset of satellite images. About half of these images (~100k in total) contain orphan wells and half do not contain orphan wells. This uses two different datasets for orphan well locations, from two different years, to identify the recently discovered ones. These images are collected from the National Agricultural Imagery Program (NAIP). Our next steps are to release the dataset to the community and train a machine learning model to segment the images into regions that contain a well and regions that do not.

LANL-Quantifying methane emission from infrared 2D images

Project objective: Develop deep learning model to quantify the methane emission from 2D images to prioritize leaks for future operations.

Summary of work: LANL developed an initial data set and ML model with synthetic 2D images that mimics Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) data. Our model is showing good performance compared to alternatives, on low emitters. Low emitters are our focus since orphan wells will typically be small leaks compared to existing machine learning approaches to this problem, which focus on relatively larger leaks (e.g., from a natural gas power plant).

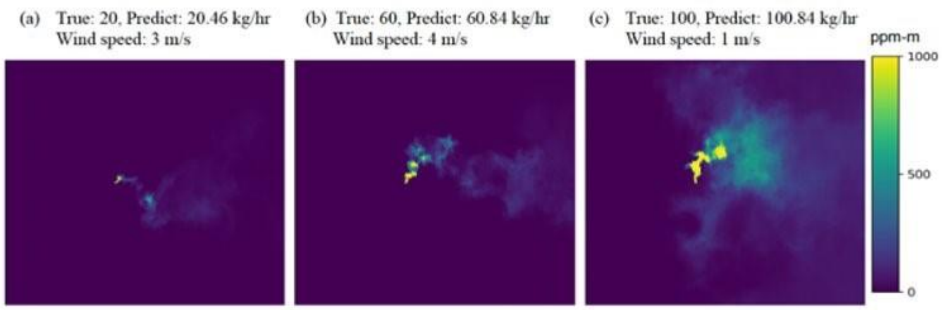


Figure 5.19: Images of methane plumes are used to estimate leak rates. The figure title compares the true leak rates with the predicted leak rates.

Next steps: Test on more infrared imaging spectrometer data, add more realistic noise, and test on regional areas with more data from orphan well which have low methane emissions.

LLNL-UOW detection and characterization using magnetic data

We are developing a sensor fusion algorithm with machine learning for UOW detection and characterization. The first problem we attack is to generate synthetic magnetic survey data to train a machine learning model for inversion of magnetic data. Disparate data (e.g., maps, methane data) will be added for sensor fusion.

A steel casing is magnetized in the geomagnetic field and induces a magnetic anomaly that can be detected by a surface or airborne magnetometer. We modeled the magnetic response of a vertical steel casing to assess detectability of a casing by magnetic surveys. A fictitious steel casing is located near Bakersfield, CA with these casing parameters: casing outer diameter of 10 inches, wall thickness of 1 cm, length of 20 m, depth of 0 (right below the surface), and susceptibility of 200 SI. A magnetometer is 2 m above the ground surface. The total magnetic anomaly is similar to a monopole response. The peak anomaly is about 1300 nanotesla (nT), much higher than the typical 1 nT noise level (Figure 5.20).

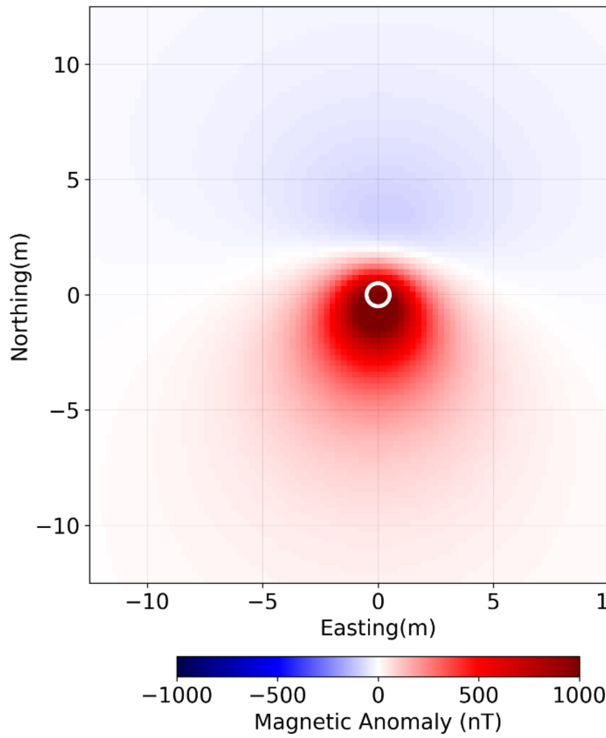


Figure 5.20: Synthetic magnetic anomaly of a steel casing. The white circle is the location of a vertical steel casing. Color indicates the intensity of the magnetic anomaly.

We studied the impact of casing and sensor parameters on the magnetic anomaly. The figure 5.21 shows (upper left) that the magnetic anomaly decreases rapidly as the magnetometer altitude increases. When the magnetometer altitude is 45m, the anomaly decreases to about 1 nT, a typical noise level. This result provides a good guideline for airborne magnetic surveys. The upper right panel shows that when the casing length approaches 20 m, the magnetic anomaly reaches a peak value. The lower two panels show that the magnetic anomaly is linearly proportional to casing diameter and susceptibility. Magnetic surveys have the potential to resolve casing diameter, wall thickness and susceptibility, but not the casing length if it is longer than 20 m. Ground magnetic surveys obtain a stronger anomaly than the drone magnetic surveys that are more efficient and cost effective to survey a large area.

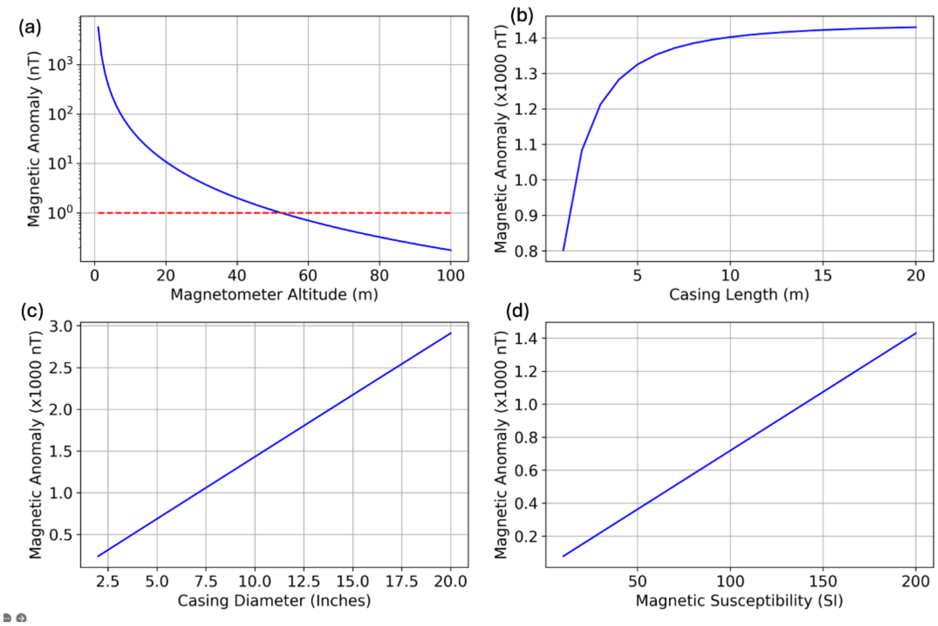


Figure 5.21: Parameters affecting magnetic anomaly of a steel casing. (a) magnetometer altitude. (b) Casing length. (c) Steel casing length. (d) Magnetic susceptibility of the steel casing.

We obtained three airborne magnetic data sets from EDX for training and testing ML models for detection and characterization of UOWs. Three airborne magnetic survey sites are located at Sanchez (Cranfield), MS, Rocky Mountain Oilfield Testing Center, and Burgettstown, PA respectively. These field magnetic data will be used for transfer learning. NETL CO2-Locate well database on EDX provides well casing information that labels the wells for supervised training.

5.5 Future Work

LLNL has received approval to conduct on-site methane releases to test different sensors. LLNL is planning on initially testing the Laser Falcon and the Aeris closed-cell methane detector. LLNL is also planning on setting up a test area on site for looking at the Magnetometer and the pulse-ekko GPR. WP3 will be using data collected at the LLNL site and field campaigns conducted by WP9 to aid with the machine learning/data science portion of WP3.

5.5.1 Selection, training, deployment, and benchmarking (ML and other) of algorithms

Apply the workflow to other producing states to estimate the total number of OWs in order to identify regions of interest. Develop algorithms to localize the regional area (Zip Code or smaller region) using GIS, location of IOWs and the estimated OWs of a county.

6 Work Package 4- Characterization

The following team members contributed to the writing of this chapter: Jaisree Iyer (LLNL), Greg Lackey (NETL), Preston Jordan (LBNL), Christina Morency (LLNL), Yuxin Wu (LBNL), GD Beskardes (SNL), CJ Weiss (SNL), Andrew Delorey (LANL).

6.1 Section Summary

Work package four (WP4) is characterization, which consists of three major items in FY23:

1. Documenting State priorities and best practices.
2. Determining the environmental hazards of orphan wells.
3. Determining the physical characteristics of orphan wells to inform the Plugging and Abandonment plans.

One multi-lab team is working on item 1, overlapping with WP 7 (Data Extraction). One lab (LANL) is working on item 2, quantifying point leaks from the well infrastructure using Gaussian plume modeling and ambient methane and wind observations, measuring the ethane in the natural gas that depends on the reservoir composition and subsurface migration processes, and developing methods to monitor diffusive leaks using a chamber. These activities overlap with other work packages, such as WP 1 (Methane Quantification) and WP 5 (Best Practices) and team members on WP4 are also working on the overlapping work packages and reporting progress to those work package teams. There are 4 different teams working on item 3, applying 4 different methods. The methods being developed and evaluated are: (1) Time Domain Reflectometry (TDF), which consists of propagating an electrical pulse down the well casing and recording various reflections, also recorded at the well casing, associated with damage or discontinuities, (2) Ground Penetrating Radar (GPR), which consists of airborne measurements that are sensitive to both cased and uncased wells, (3) Acoustic characterization, which consists of recording traveling waves in the well casing due to ambient noise and using travel times to determine the depth of the well, and (4) Direct Current Electrification (DCE), which consists of applying a current at the well casing and measuring the electrical field surrounding the well either at the surface or in a nearby well, revealing depth and damage information about the well.

Of these 4 methods, 3 require direct access to the well head or well casing, while the GPR method makes airborne measurements in the local vicinity of one or more wells and in fact can potentially detect unknown wells (a goal of the overall program, but not of this work package). Each method has some potential strengths and weaknesses both in terms of what information about the well they can provide, and the cost and complexity of the measurement themselves. We are deliberately focusing on methods that do not require putting a tool or sensor into the well, as that adds cost and complexity. We may end up recommending simple down well techniques like lowering a camera or some other low-tech sensor, but these options are not currently being investigated. We understand that a suite of measurements that can be performed simultaneously or synergistically, within this work package and with other work packages, will likely be the most effective way to characterize a well for plugging and abandonment. In the next section, each team has written a description of their approach and progress.

6.2 Activity

6.2.1 Documenting State priorities and best practices

Well scoring systems were requested from 22 states agencies and the U.S. Department of Interior (DOI). Responses from 15 state agencies and the U.S. DOI provided a well scoring system with sufficient detail to perform the analysis. Seven narratives describing notional example wells were generated by the team and scored blindly by another team member using each state scoring system.

A preliminary analysis of the state scoring protocol was completed in September 2023 and presented to members of the Interstate Oil and Gas Compact Commission (IOGCC). Synthesis of the orphaned well scoring protocols found general alignment between states. The top three criteria prioritized by states were: 1) distance to sensitive receptors, 2) leaking status, and 3) safety or public concern. Scoring of the notional example wells also aligned between states with higher risk wells prioritized over lower risk wells. One exception to this trend was the scoring of the well narrative that described an undocumented well. This well received a top-three score by six of the 12 scoring systems compared in the analysis. However, it also received a bottom-three score in the other six systems highlighting discrepancies in the way states currently account for missing information in their scoring systems. This work clearly identifies and summarizes state priorities for well plugging. The CATALOG team will leverage this work to focus their undocumented orphaned well characterization efforts on the identified priorities. This effort also creates a resource for state regulatory agencies. We benchmark the considered orphaned well scoring systems with a standard risk assessment protocol and compile a synthesized dataset of the scoring systems used by each state. A study draft describing the progress made thus far on this effort is being prepared. Future work will entail comparing orphaned well scoring approaches using real state scoring data, gathering plugging costs and cross-referencing them with well scoring information, performing a state-level analysis with available well information, and supporting the development of the optimization tool in WP 4.

6.2.2 Determining the physical characteristics of orphan wells to inform the Plugging and Abandonment plans

LBL Time Domain Reflectometry: LBL is developing a novel borehole characterization technology based on electromagnetic time domain reflectometry, EM-TDR. Our EM-TDR method is based on guided wave principles (Figure 6.1) and offers a few key advantages that are discussed below. EM-TDR aims to provide ultrafast diagnosis of borehole health conditions by measuring EM signals propagating along the boreholes with sensitivity to corrosion or stress related damages, and other features, such as bottom of the casing. Briefly, EM-TDR utilizes the steel pipe itself as a waveguide to carry short-pulsed EM waves without the need to install sensing cables. When damage induced impedance changes are encountered during wave propagation, a reflected EM signal is returned to the source end, carrying information about the location and magnitude of the anomaly (Wang and Wu, 2020). Our proposed EM-TDR technology is uniquely suited for fast screening of a large number of UOW boreholes to assess their conditions to guide more detailed characterization efforts to support plug and abandonment (PA) efforts.

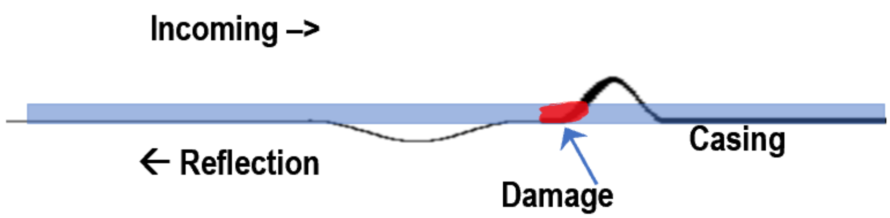


Figure 6.1: The operating principle of Electromagnetic Time Domain Reflectometry (EM-TDR) under development for WP4.

<u>EM-TDR offers a few key advantages when compared to other methods. These include:</u>

<u>Extended length of detection</u>: EM-TDR method offers an extended detection length beyond what is achievable with other testing methods due to limited signal attenuation and the lightspeed of wave velocity. The technology has been demonstrated on metallic pipes at the hundreds meter scale under sub-optimal conditions with significant signal loss and interfering signals (Wang and Wu, 2020). Under typical borehole conditions, a much longer inspection length is expected. Some recent tests on O&G wells in Oklahoma and preliminary results are shown in later sections.

<u>A new diagnostic modality</u>: EM-TDR is based on the transmission of EM signals in conductive metals, and is interrogating changes of the electrical properties of the pipe instead of the acoustic properties with UT/EMAT type of technologies. This offers a new diagnostic modality that can pick up pipeline degradation processes that produce large electrical anomalies but not acoustic. It thus offers a complementary mode of inspection. Some examples of lab testing, validating the sensitivity of the EM signal to damages, even small ones, are shown in the next section.

<u>Fast and easy deployment</u>: EM signals travel at the speed of light, which is a few orders of magnitude faster than mechanical waves (e.g., acoustic waves). A roundtrip signal takes less than a second in almost any length of borehole and is capable of interrogating a long pipe. In addition, coupling of the EM-TDR system with a borehole is simple. For the acoustic method, careful preparation of the exposed pipes is required to ensure good mechanical coupling which can be challenging with odd shaped or rough and irregular contact points. EM-TDR only requires a conductive connection with the pipe which is much easier to establish.

<u>FY23 major activities</u>

1. *System prototyping and optimization*

   Prototyping and optimization of the EM-TDR system was carried out in the past year with a focus on designing a system that will work on wellbores in legacy O&G fields. A schematic design, including the major pieces of electrical components, is shown in Figure 6.2 below. Tests conducted suggested that having a return path, often a metal conduit, can help improve the signal to noise ratio (SNR), but this is not always the case and is not required if such a return path does not exist. In a typical survey, the primary signal will be applied to the casing layer of interest but a different casing layer can be used as the return path. It may be the case that multiple layers of casing exist for a typical well, including conductor, surface, intermediate and production casing, and any of these casing strings can be inspected as the primary target. This is another unique feature of EM-TDR, where the outer layers of casings are often not accessible with borehole logging tools.

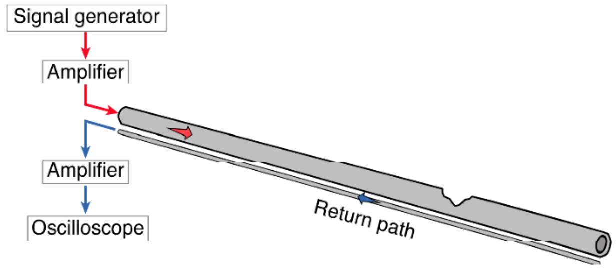


Figure 6.2: The conceptual setup of the EM-TDR system showing the major electrical components involved and the signal transmission pathway.

A key technical aspect of the EM-TDR tool is the stability of the signal entry contact on the target casing, in order to maximize transmission and minimize entry reflection. The contact point also needs to be stable enough to prevent aliasing signals or drifting of the signal during data acquisition. Initial testing at both lab and field scales suggested that the current approach of direct contact utilizing conductive tape coupling results in acceptable signals, but there is a need for continued improvement. Some of the initial lab and field-testing results are shown in the sections below.

2. *Laboratory testing of EM-TDR*

Extensive testing of the EM-TDR prototype was carried out in the last year utilizing controlled laboratory experiments where metallic tubing of different sizes with engineered damages were utilized. An example of the results is shown in Figure 6.3 below.

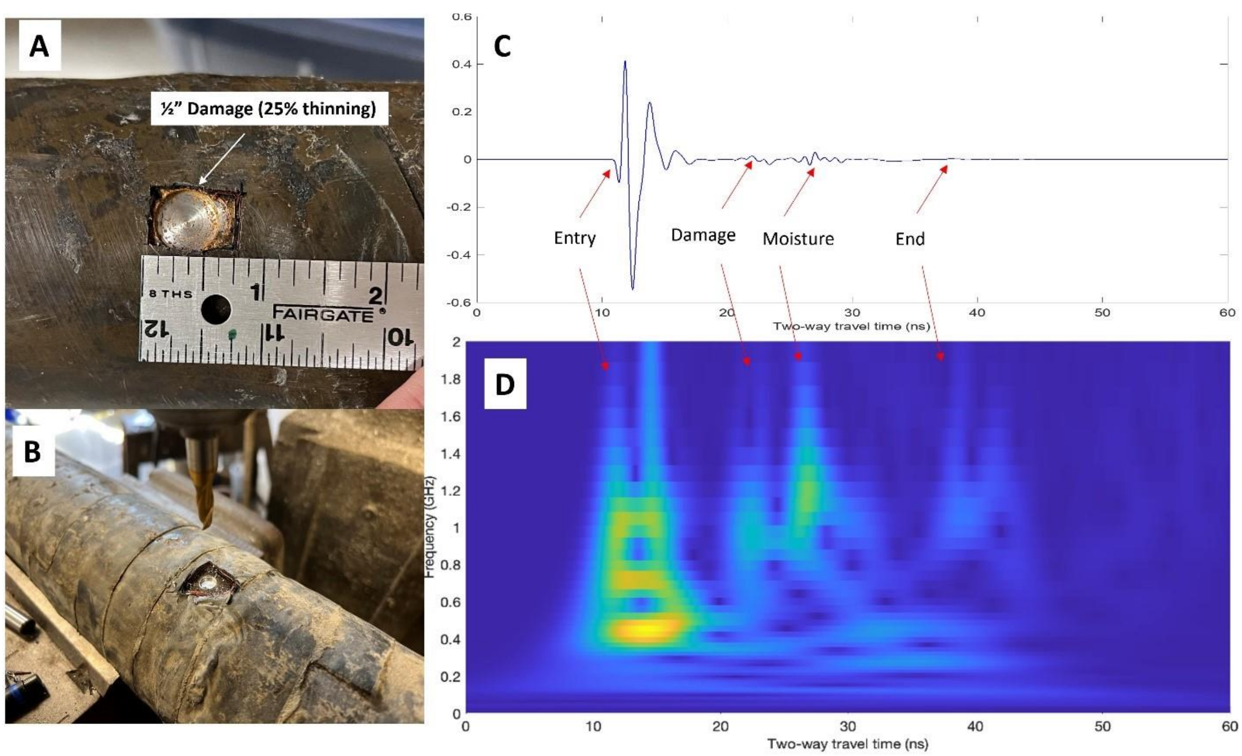


Figure 6.3: Example results of the EM-TDR tests on tubing with engineered damage. A-B: Engineered damaged made on metallic tubing acquired from an operator with a circular thinning on the tubing at 0.5-inch diameter and 25% wall thickness loss; C: EM-TDR trace for the 2GHz signal indicating the key features corresponding to those on the tubing; D: Wavelet transformed EM-TDR signals across the frequency bandwidth tested from 100M Hz to 2 GHz.

This example dataset was acquired for a piece of 4” tubing acquired from a partner facility operator. Engineered damage of 0.5” in diameter and 25% wall thinning was created in the middle of the tubing and the tubing was half buried in soil with added moisture to assess the impact of changing geology (soil in this case) and moisture content on the EM-TDR signals. Figure 6.3A/B shows the damage on the tubing before soil burial. Figure 6.3C shows the identification of the main features for the 2 GHz signal after data processing with empirical mode decomposition (EMD), which decomposes the signal into intrinsic mode functions (IMFs) that capture different frequency components of the signal. The engineered damage on the tubing is clearly visible and matches with its physical location on the tubing. In addition to the damage, the added moisture is also visible behind the damage feature. The end of the pipe can also be identified while the reflection is small when compared to other features. To improve the visibility of the features and also demonstrate the frequency dependent behaviors of the features, wavelet transform (WT) was conducted on the signal as shown in Figure 6.3D. In addition to the features identified in Figure 6.3C, Figure 6.3D also shows the expression of the features for different frequencies, where the lower frequencies (<1G Hz) present challenges in resolving the small features on the tubing with convoluted, or “smeared”, signals blending multiple features together. On the other hand, those frequencies >1 GHz are able to identify the features., The frequency band between 1.2 – 1.5GHz appears to be optimal for this particular test, and is able to resolve the features while maintaining the higher magnitude of the signals. Fundamentally, there is a tradeoff among signal resolution, magnitude and depth of penetration that needs to be optimized for a given testing scenario.

3. Numerical modeling and simulation of EM-TDR signals

The numerical simulations conducted are aimed at helping to understand the EM-TDR behaviors for borehole casing and tubing, and to understand potential signals associated with borehole damages, such as those induced by corrosion. As a novel technology, EM-TDR is still at an early stage and the numerical simulations conducted here are critical for sensitivity analysis. The simulations discussed here are using a 2.5-D simulation procedure instead of a full 3D simulation. Due to the extreme ratio between the casing length and wall thickness, meshing and computation of a full 3D model is very challenging to run. In the 2.5D approach, first we simulate the cross-section of the model to compute the impedance properties by solving Maxwell’s equations and the associated electric field and current distribution. Then, using the cross-section of the model, we calculate the unit length characteristic impedance. Finally, we assemble the characteristic impedance from different sections into a longitudinal 1-D simulation. The detailed background of this method can be found in Wang and Wu (2020).

A typical O&G well setup is used to simulate the EM-TDR responses to impedance changes. An example of the 2D simulation with corrosion features is shown in the figure below (Figure 6.4).

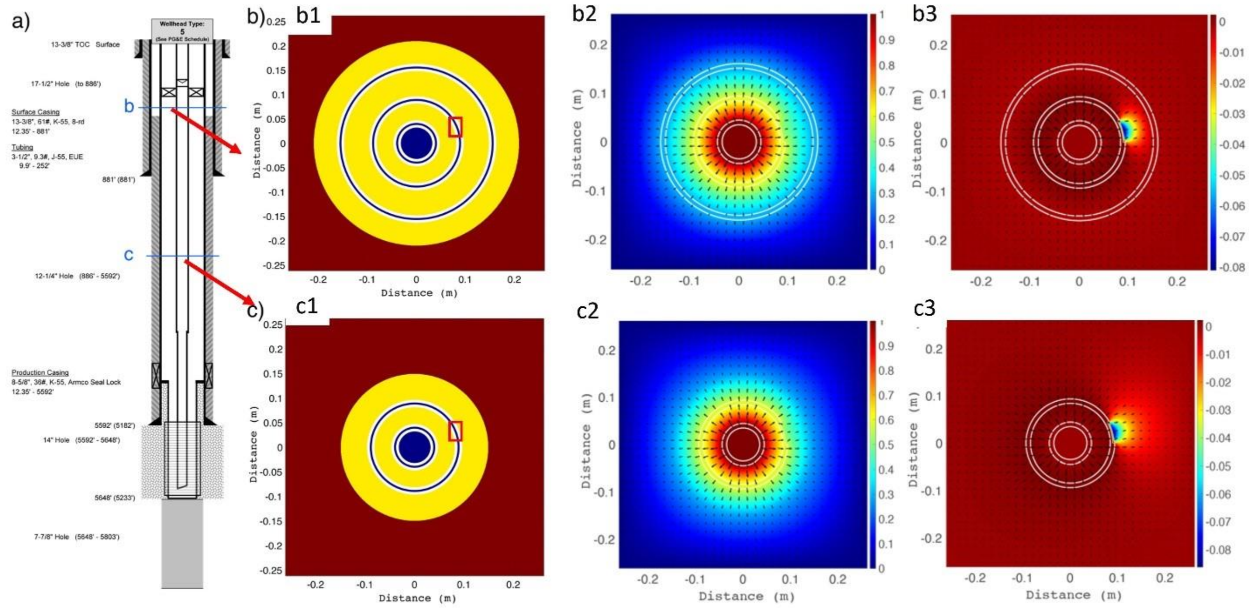


Figure 6.4: Numerical modeling and simulation of the EM fields for O&G wells due to corrosion damages. (a) The model wellbore diagram used in the simulation where two cross sections (b and c) are used for the 2D simulation of the EM responses due to corrasion. The top row of colored b figures (b1-3, c1-3) represents the 3-casing layer borehole cross section showing the corrosion damage feature in the red box (b1), the potential field distribution when the casing is energized (b2), and the changes in the potential field due to the corrosion damage (b3). The bottom c row shows the same features for the damage at a deeper location (c) shown on the borehole diagram where only 2 layers of casing are present.

The simulations shown in Figure 6.4 aim to evaluate the magnitude of the EM anomalies due to damage features, such as casing corrosion, in order to help evaluate the sensitivity of the EM-TDR signals to this damage. Many more simulations were conducted representing damage on different strings of casing and of different magnitudes, as well as different fluid compositions in the borehole (e.g. oil vs water) and the invasion of fluids into the damaged areas. These 2D simulations are assembled into 2.5D simulations based on the approach described at the beginning of this section to help produce the forward simulation of EM-TDR responses that are important for the eventual numerical inversion and interpretation of the EM-TDR signals from the actual datasets.

4. *Initial field testing in Oklahoma*

Initial field testing of the EM-TDR technology was conducted in Oklahoma on 5 different wells located in Holdenville (3 wells) and the University of Tulsa campus (2 wells), as shown in Figure 6.5 below.

Figure 6.5: O&G wells in Oklahoma used for the initial field testing of the prototype EM-TDR system. Wells # 1-3 are located in Holdenville, OK and Wells # 4-5 are located on the University of Tulsa campus. Wells #1-4 are 5 ½ inch diameter wellbore wells while Well #5 is a 12 wellbore well. All wells are at the depths ~ 3000 ft.

Interpretation of the field-testing results are ongoing and a preliminary analysis from one of the wells (Well #1) is shown in Figure 6.6 below.

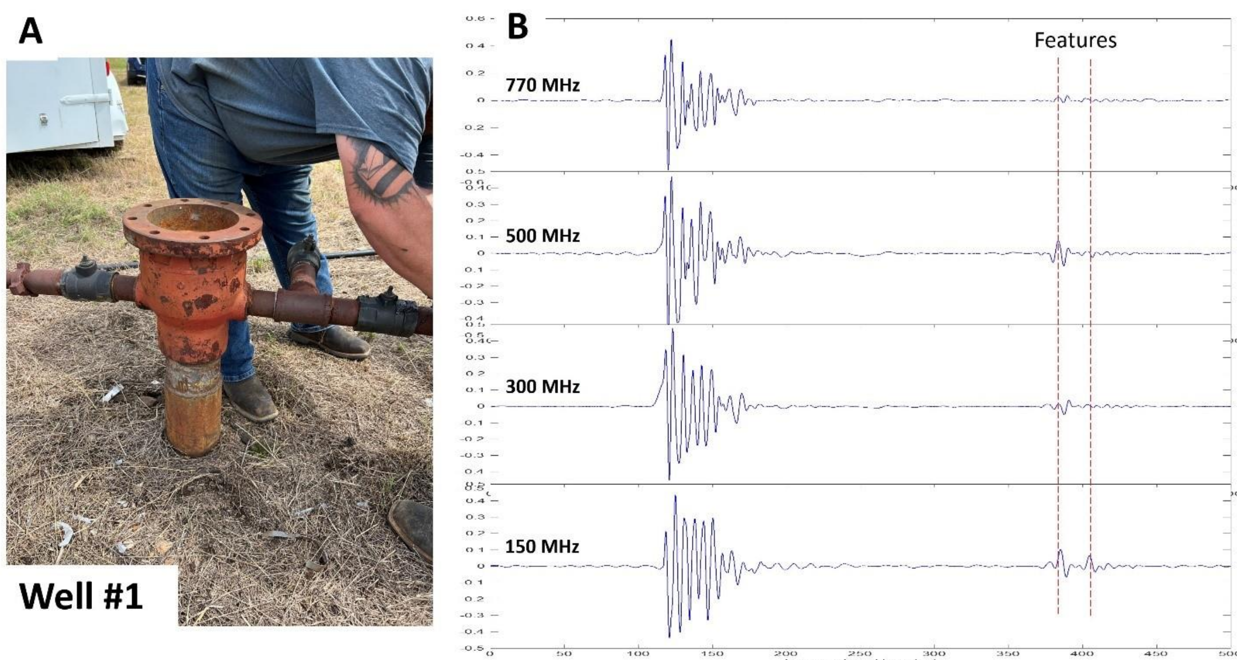


Figure 6.6: Example EM-TDR dataset from an O&G wellbore in Holdenville, OK. A: The test borehole; B: EM-TDR traces for frequencies at 770 MHz, 500 MHz, 300 MHz and 150 MHz. The identified features are marked as dashed red lines.

Figure 6.6 shows a preliminary example of the EM-TDR data for the first 500 nanoseconds (ns) of the wave traces. Consistent features are present across all frequencies around ~400 ns (~ 30-40 meters into the borehole), representing a high likelihood of anomalies present on the casing. A total trace of 5000ns is recorded for each dataset, and these are being processed currently.

The nature of these anomalies needs to be confirmed and validated using traditional borehole logging techniques, which will be planned in collaboration with the borehole owner, a local independent operator, and a borehole logging company interested in collaborating with us. Note that EM-TDR is sensitive to the entire cross section of the borehole, including behind casing features, such as cement and bonding, which might be difficult to identify through visual borehole logging, if the damage feature does not present on the inner surface of the casing.

Summary of accomplishments

The main accomplishments of at LBNL this task can be summarized below:

- A prototype EM-TDR system was designed and constructed for borehole integrity characterization
- Numerical simulation capacities were established to simulate EM responses to borehole damage, such as corrosion, using a 2.5D modeling approach
- EM-TDR signal processing algorithms using EMD and WT were established to improve feature identification and visualization
- Laboratory testing successfully validated the performance of EM-TDR in identifying damage features on borehole tubing.
- Initial field testing of the system on multiple O&G wells was conducted with preliminary data showing potential features pending validation.

6.2.3 Ground Penetrating Radar (Priority 3, LLNL Lead)

(LLNL) Numerical development:

During this first year, we developed and adapted our in-house EM spectral-element code, SPECFEM_EM, to mimic airborne GPR acquisition and set the stage for subsurface imaging serving both well detection, especially those targeting uncased wells (WP2), and well characterization (WP4), through inverse approach based on adjoint method (Figure 6.7).

Preliminary results revealed the following findings. Focusing first on a 2D geometry, we designed a mesh with 15 m of air layer on top of a 15 m deep subsurface with a 15 cm diameter uncased well 3 m deep. To mimic the drone movement, we consider 11 source-antenna pair locations at 1 m above ground over a distance of 10 m centered above the well location (Figure 6.8).

Figure 6.9 shows a snapshot of the electric field at 17.5 ns, with and without the presence of a well. While the signal is strongly dominated by the direct and surface reflected waves, we do observe clear signals associated to the well, such as transmitted waves, head waves, and reflected waves. We are most interested in the reflected waves recorded by the drone, which, with further analysis, can help recover the well location. Radargrams confirm the dominance of the direct and surface reflection waves but removing them highlights the presence of the well signature signals (Figure 6.10).

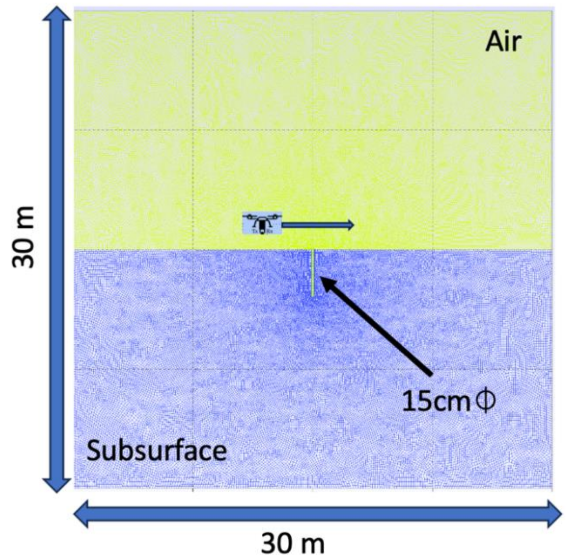


Figure 6.7: 2d mesh used to mimic airborne GPR acquisition. The mesh contains 83,812 unstructured quadrilateral spectral-elements. GPR source is a Ricker function of central frequency 500 MHz. A schematic drone is displayed showing altitude of 1 m and direction of the acquisition.

Figure 6.8: Zoom in showing the location of the 11 source- antenna pair locations with respect to the well. Sources are displayed as yellow crosses and receiver antenna as green squares.

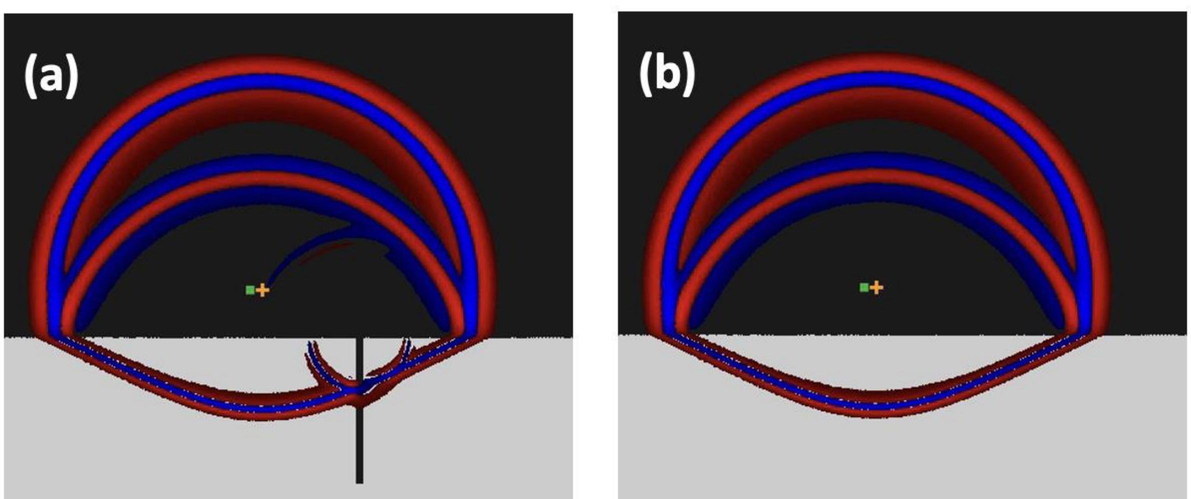


Figure 6.9: Zoom in snapshots of the electric field at 17.5 ns for source-antenna pair at location 4, offset from the well location by 2 m. Panel (a) depicts the wavefield with the presence of a well, while panel (b) depicts the wavefield without a well.

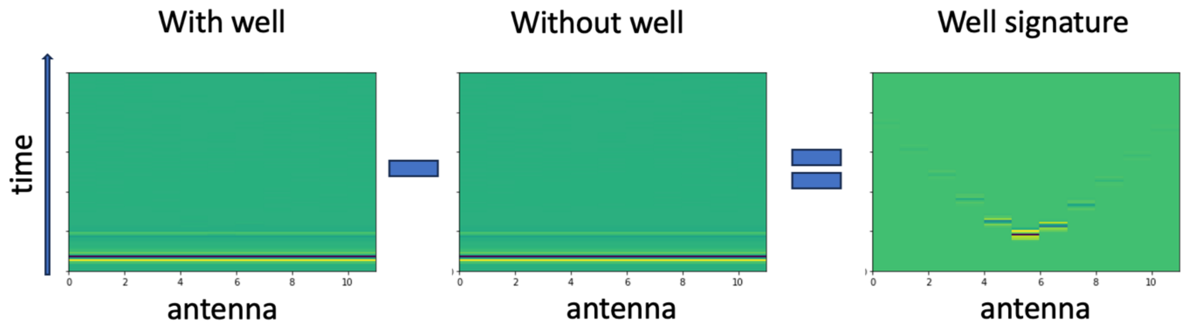


Figure 6.10: Stacked radargrams with and without well, and their difference. Raw radargrams present two strong signals: direct and surface reflected. The difference highlights the presence of the well in the classical form of a parabolic signal.

LANL Acoustic Measurements

(LANL) We are testing acoustic methods for determining the depth of a well. Using a well on the LANL campus, we performed a few simple tests to determine the viability of different acoustic signals. This well is primarily dry and so we don't expect acoustic behavior associated with a fluid-filled pipe. It has a single casing constructed of stainless steel, has a depth of 331m and contains an accelerometer at the bottom (Figure 6.11). The accelerometer 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 a piezoelectric transducer at the top of the casing as we would in practice to compare the results. Initially, we are testing the transmission of sources through the casing as well as the ambient conditions. To test the transmission of a source through the casing, we struck the top of the casing with a hammer to see if a signal was recorded at the bottom of the well. In a real scenario, we would have a sensor at the top and would be looking to record a round-trip elastic wave. A hammer blow is impulsive and can generate a broadband signal, but in practice we can produce frequencies up to hundreds of Hz. We were not able to detect a signal at the bottom of the well associated with our hammer blows and so we do not expect this to be a viable signal for deeper wells and for round-trip signals since we were not able to detect a one-way traverse of the casing. Using a more energetic source would risk damaging the well head.

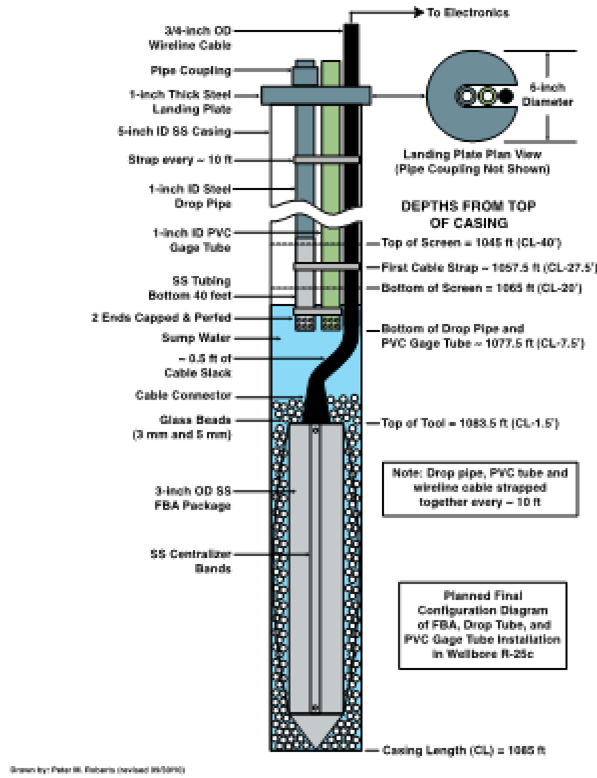


Figure 6.11: Diagram of the well we used to test acoustic methods.

The second test 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, moving vehicles, nearby operating 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 very weak, but by recording over time, we can extract and amplify weak signals through stacking.

We have several years of data available from the accelerometer but used only about two months' worth. 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 60Hz due to the frequency of alternating current. 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, that we know the path of the wave, and that we know 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 is not related to the well casing and finally there could be standing waves in the casing. Some of these possibilities could be resolved with finite element modeling, which we have not yet performed at this time. With a little bit of prior information, it is possible to isolate the signal that we want. As an example, if a signal repeats at a frequency (*f*) the one-way distance traveled would be d = (v/f)/2, where *v* is the assumed velocity of the wave. In this case, we know the P-wave and S-wave velocity of stainless steel, and can estimate the velocities of the interface waves, though they depend on knowing the fluid and formation velocities. We can also estimate the P-wave and S-wave velocity strictly in the geologic formation. We can examine all the repeating signals in the auto-correlation and eliminate those that cannot be a wave traveling through entire length of the casing. The remaining signals can be considered for estimating the length of the casing.

In Figure 6.12, we show the spectra for the auto-correlation, but convert the horizontal axis to velocity, if the distance traveled is 662m, which is double the length of the casing and the travel distance of a trapped traveling wave for each cycle. We can see that there is a peak exactly at the shear wave velocity of the casing, suggesting that we could use shear wave velocity to determine the depth of the casing. According to these results, the shear wave appears mostly on the east component, and secondarily on the vertical component, but not on the north component. If this is a shear wave, it is highly polarized. There are two other peaks, the first of which could be a wave bouncing around in the formation rather than in the pipe; it is too slow to be a wave in the casing. The peak at the right could be a reflector closer to the sensor as it appears much too fast to be a wave traveling the entire casing. Perhaps due to the well having no fluid in it, the various tube waves are not observed here.

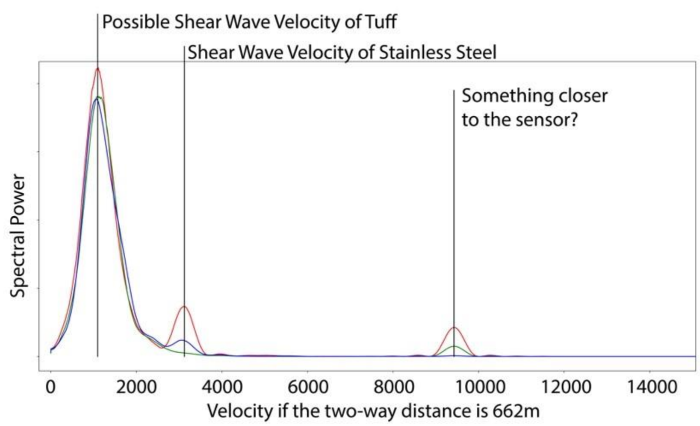


Figure 6.12: Spectra for auto-correlation with horizontal axis converted to velocity if the travel distance is the two-way distance through the well casing. Red, green, blue indicate east, north, vertical, respectively.

6.2.4 Rapid, scalable 3D Computational Modeling (Priorities 2-3 SNL Lead)

SNL Top-Casing DC Electrification

Quantification of wellbore integrity is one of the key characterization objectives to inform prioritization and risk assessment for subsequent repurposing/plugging of newly documented orphaned wells. Here we present a numerical study on the feasibility of direct current (DC) electrical methods in this effort (Beskardes and Weiss, 2022). The survey setting includes a steel-cased well whose condition is unknown (e.g., intact or damaged) and a radially-emanating surface profile of electric field measurements arising from a top-casing electrification of the well (Figure 6.13). Damage along the well-bore results in localized changes in its electrical conductivity, which are then reflected in both the amplitude and character of electric field along the profile (Figures 6.14 and 6.15).

DC electric field estimates are computed efficiently using a hierarchical finite element method (Weiss, 2017) whereby the electrical properties of wellbore are associated with the edges of an unstructured tetrahedral mesh. This strategy avoids the costly problem of direct discretization of the wellbore itself, thereby allowing for rapid computation of modest computational resources – in this case, a MacBook Pro laptop. It also paves the way for direct inversion of electric field data on commodity computers as the cost of the forward problem (now, drastically minimized) is the singular sticking point in computing iterative model updates during inversion.

Using a Levenberg-Marquardt least squares methodology for computing model updates, we demonstrate that constrained inversions – where the location of the damage zone is presumed *a priori* – are able to recover wellbore conductivity values acceptably close to the known values using surface-based data alone (Figure 6.16). Recovery of the damage zone location, however, is poorly constrained by these data. To overcome this shortcoming, we show that a combination of surface data and data obtained by a nearby, uncased "observation" well provide sufficient information to also constrain the damage zone location (Figure 6.17).

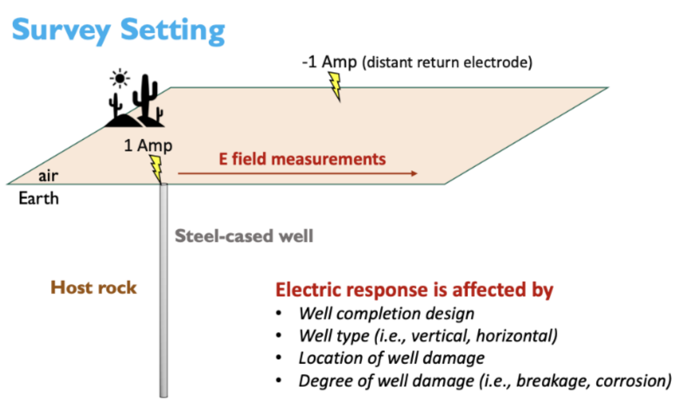


Figure 6.13: Modeled experimental setup for "top-casing" method of wellbore integrity assessment (Wilt, 2016; MacLennan et al., 2018).

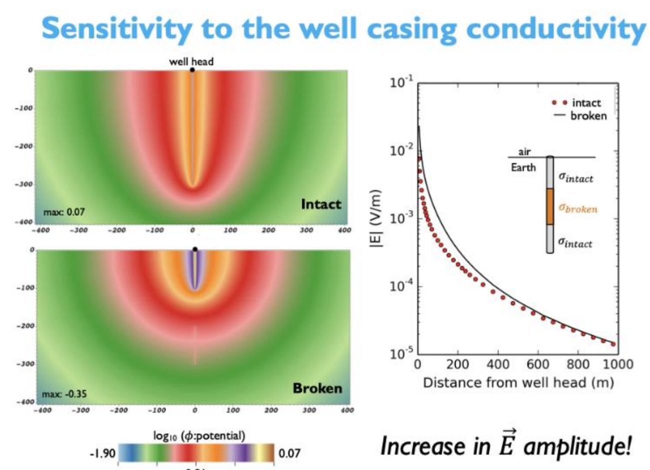


Figure 6.14: Distribution of electrostatic potential in the subsurface (left) arising from top-casing excitation for a fully in-tact well (left, top) with uniform electrical conductivity, and a well with a mid-depth region of reduced conductivity due to

"damage" from corrosion, breakage, etc. Surface profiles of radial electric field (right) show that the damaged well is distinguishable from the intact well through differences in its near-offset curvature and amplitude.

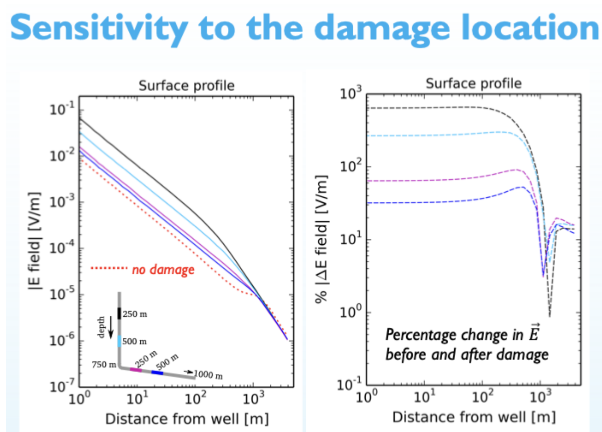


Figure 6.15: Profiles of radial electric field (total field, left; relative field, right) on Earth's surface over the horizontal section of a deviated well show the effect of the damage location on measured response *(Beskardes et al., 2021).*

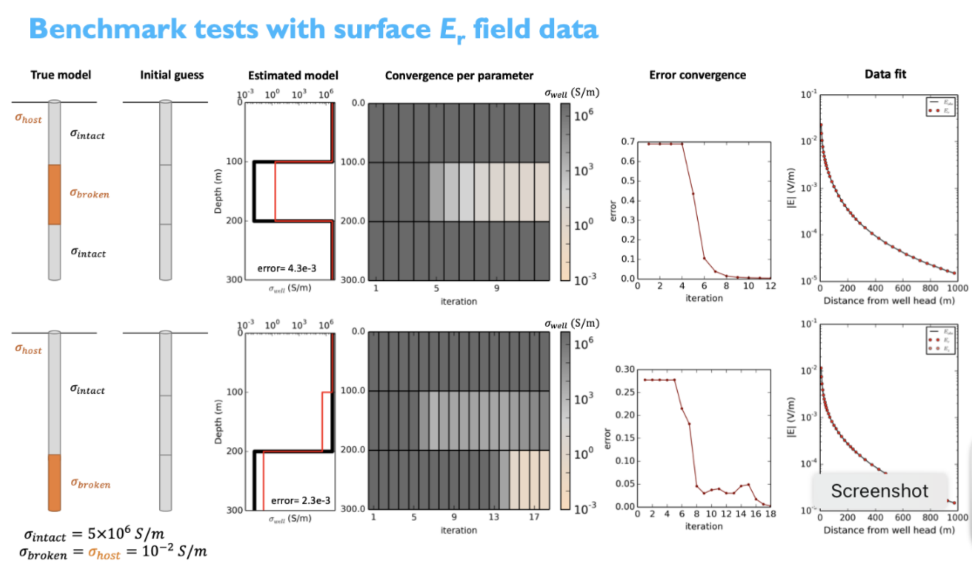


Figure 6.16: Inversion results for damage zone quantification using surface-based E-field measurements and top-casing excitation for mid-well (top row) and bottom-well (bottom row) damage scenarios. Constraining the free parameters in the inversion (4th column) to the regions where the damage is believed to lie allows for acceptable recovery of the damage zone conductivity (3rd, 5th and 6th columns).

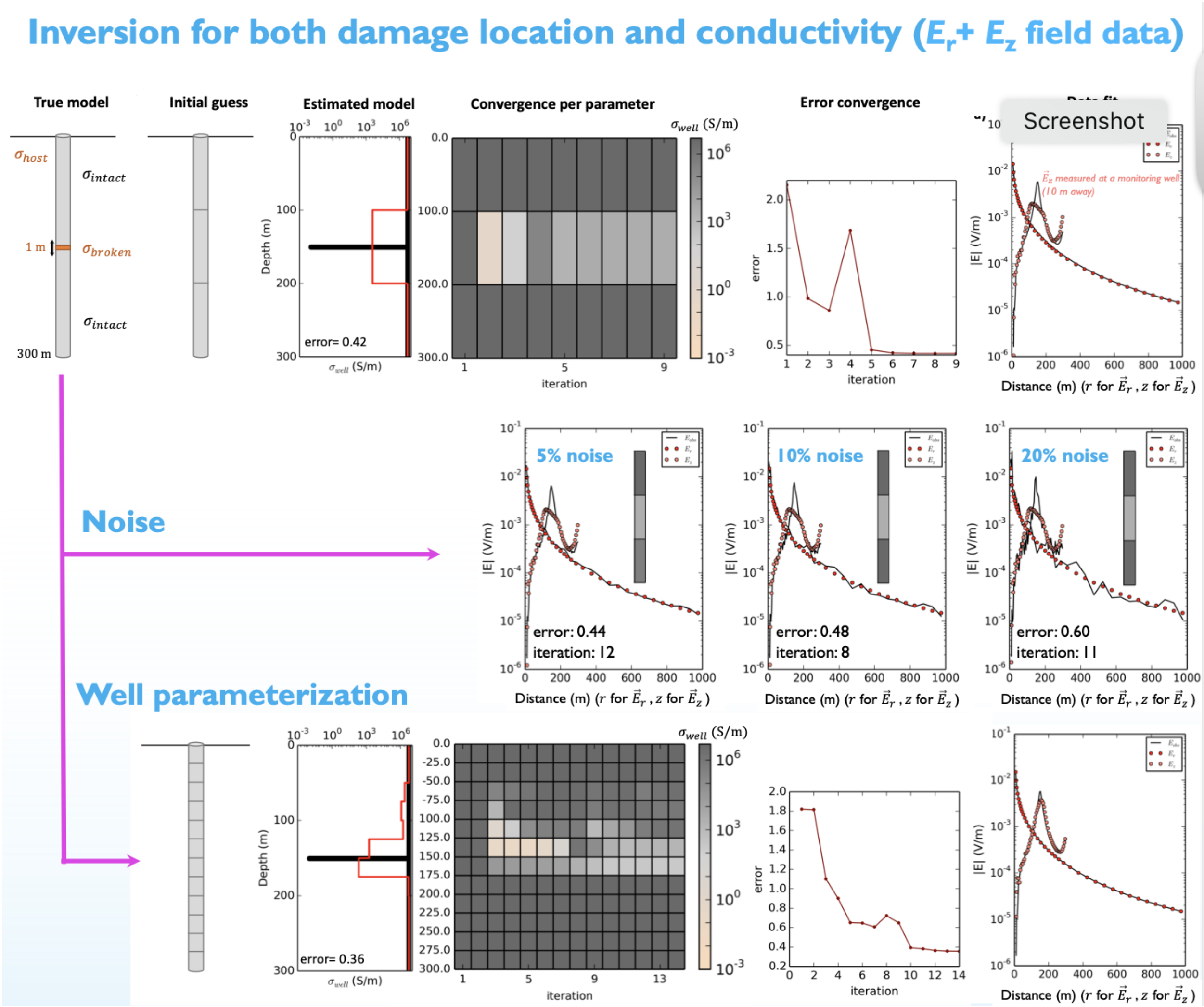


Figure 6.17: Inversion results for constraining both well damage magnitude and location using a combination of data collected on both the surface and inside a neighboring "observation well". Damage location is reasonably constrained, regardless of discretization coarseness (top and bottom rows) and robust with respect to bounding estimates of observational noise (middle row).

6.3 Motivation and Background

As of April 2022, 123,318 orphaned wells were reported by state oil and gas regulatory agencies and as such were eligible for federal remediation funding through the Bipartisan Infrastructure Law (BIL) (Interstate Oil and Gas Compact Commission (IOGCC) 2021a; Boutot et al. 2022). This was approximately a 50% increase from the 81,857 documented orphaned wells reported in September 2021. Additionally, it is estimated that there are between 310,000 to 800,000 undocumented orphaned wells that will need to be located and characterized prior to plugging (Interstate Oil and Gas Compact Commission (IOGCC, 2021a). Plugging costs for the ~130,000 documented orphaned wells are estimated to be between \$6.3 and \$8.4 billion, which far exceeds the \$4.7 billion made available for well plugging by the BIL (Kang et al. 2023). This highlights the need to focus orphaned well plugging efforts on the highest priority wells that pose the greatest risks to human health and the environment.

State oil and gas regulatory agencies were required to provide a description of their prioritization processes as a requirement for BIL grant applications (IOGCC, 2023b). The plugging prioritization processes submitted by states can be generally described as a system for assigning scores to wells based on their attributes. Scores for well attributes vary and are allocated based on the unique priorities of each state. For example, a state may assign a higher score to wells that are actively leaking and a lower score to wells that are in good condition. Wells with the highest scores are targeted for plugging first. Analyzing the scores assigned by states to each well attribute provides insight into their priorities. We synthesized and compared well scoring systems for 15 state agencies and the U.S. Department of Interior to identify well attributes that are universally important for well plugging. Our findings inform and guide the characterization efforts of the CATALOG well characterization efforts under Work Package 4.

Abandoned and left without proper documentation, undocumented orphaned wells (UOWs) present challenges to stakeholders seeking to locate and access them. The limited availability of advanced imaging and characterization technology tailored for assessing their borehole conditions, as well as their potential environmental impacts, further complicate assessment, and remediation processes. Additionally, the high cost associated with deploying wireline tools in such wells, traditionally used for well diagnostics and evaluation, can be prohibitive. Yet, there is an urgent need to characterize these wells quickly and cost-effectively. Rapid and affordable characterization is necessary to support timely plugging and abandonment efforts, ensuring that potential environmental risks are mitigated before they escalate. This urgency adds another layer of complexity to the already challenging endeavor of managing orphaned wells.

6.4 Data Summary

TDR data is collected by LBNL in laboratory tests and in the field in Oklahoma. Seismic/acoustic data is collected by LANL on a well on the LANL campus. Data simulations were performed at LBNL (TDR), LLNL (GPR), and SNL (DCE).

6.5 Future Work

(LBNL) Major activities for the next year will include the following:

- Complete data analysis from the first field trial and discuss with borehole owner and logging partners to plan for validation efforts
- Improve EM-TDR design and optimization based on field testing results and perform additional controlled lab tests
- Develop and improve the signal processing algorithms to enhance feature identification, SNR and visualization
- Work with stakeholders, e.g., Osage Nation, to identify new wells for field testing the EM-TDR system, ideally with known borehole conditions or parallel downhole logging efforts
- Explore joint testing opportunities with other characterization tools, such as acoustics or ground penetrating radar, being developed in parallel at other Labs.

At LLNL, future work will consist of reproducing experimental dataset to be collected at LLNL, and adapt the adjoint imaging and tomography technique to airborne GPR.

At LANL, future work will include continuation of acoustic tests and the application of different kinds of sensors in different configurations and collection of data at one or more additional wells. Numerical modeling will be performed to help interpret the observations.

7 Work Package 5- Integration and Best Practices

Integration and UOW Well Finding and Characterization Best Practices

7.1 Section Summary

An efficient transfer of knowledge gained as part of the CATALOG program is needed to enable the transition of research & development results into practice. The primary challenges associated with knowledge transfer include standardization and integration of data needed for application of UOW identification and monitoring tools, the verification of results from field testing, and the description of best practices for locating and characterizing UOWs on a region-specific basis.

7.2 Activity

Development and testing of a methodology for quick and inexpensive methane emissions estimates from orphaned wells was initiated. The methodology utilizes downwind concentration measurements and wind field measurements, coupled with a plume dispersion model, to estimate methane emissions. Although the methodology does not provide the same high level of accuracy as flux chambers or high flow sampling, it allows for easy identification and the general classification of wells as “not leaking”, “leaking at low or mid-rates”, or “high emitters that may warrant additional investigation”. With this methodology, wells that should be high priority for plugging due to high methane emissions can be quickly and inexpensively identified.

The CATALOG team has reported a methodology for systematically compiling publicly available digital data prior to embarking on field campaigns. Geographical Information System (GIS) technology is used to analyze digital data and create maps to guide field work activities based on the goals of a specific project. This workflow was developed in the Appalachian region and although certain aspects may be unique, the general process should be applicable to locating undocumented wells in other regions. The report of the methodology was designed to assist states, particularly in the Appalachian Basin, in their initial investigations of well locations for an identified area of interest.

The NETL team completed a report of best practices for conducting aerial magnetic surveys for well finding. The sharing of the report will allow stakeholders to benefit from the decade of experience that NETL has in testing and demonstrating this technique.

For field testing and demonstration of technologies that are developed under the CATALOG program, the WP5 team has focused on: (1) assigning coverage of sites across the U.S. such that technologies and methodologies are relevant and successful in a variety of settings (across differing geology, topography, land use and cover, etc.), and (2) matching the needs of specific technology field tests and demonstrations with an available site.

WP5 created a glossary ("Key Terminology for the CATALOG Project") summarizing key, relevant terms and associated generalized definitions – noting certain synonymous terms parsed out by state – to help streamline discussions and ensure consistent understanding across the project.

7.3 Motivation and Background

States and provinces have varying approaches and resources for managing the risks from orphaned oil and gas wells. Available funds are insufficient to mitigate (by plugging) all known orphaned wells across the U.S., not including undocumented orphaned wells. As such, state and tribal agencies need efficient strategies to find orphaned wells, collect and document adequate information to characterize them, and prioritize plugging operations to maximize reductions in environmental impacts. In Spring 2022, the CATALOG team received responses from a questionnaire distributed to state representatives from the Interstate Oil & Gas Compact Commission (IOGCC). The questionnaire was designed to determine what factors are important to consider, and what the states' priorities are, with respect to orphaned wells. Results are described in O'Malley et al. (2023), but in summary, some of the biggest data needs/gaps were identified as location, ownership, well construction, mechanical integrity, surface impacts, and methane emissions. Factors that states consider in ranking orphaned wells for priority plugging include leaking fluids, proximity to people and water, age and construction, mechanical integrity, wellhead pressure, hazards to wildlife and navigation, and economic efficiencies. There are clear opportunities for the CATALOG team to provide tools and technologies to fill in the knowledge gaps that can hinder the development of more effective plugging prioritization strategies.

An efficient transfer of the knowledge gained through the CATALOG program is necessary to enable the transition of research & development results into applied practices. The primary challenges associated with knowledge transfer include standardization and integration of data needed for application of UOW identification and monitoring tools, verification of results from field testing, and compiling descriptions of best practices for locating and characterizing UOWs on a region-specific basis. The objectives of the activity under this work package are to provide integrated and standardized data and to develop best practices recommendations for well finding, emissions measurements and characterization methodologies for use by Federal, State, and Tribal agencies.

7.3.1 Best practices: Use of downwind methane concentrations measurements and wind field measurements to model and estimate orphaned well methane emissions

As discovered during the IOGCC survey as well as from personal conversations between the CATALOG team and representatives from several state agencies, one of the biggest challenges to characterizing orphaned wells is determinizing how best to measure their emissions. However, there is no standardized methodology, and approaches that have been reported in the literature vary in performance metrics such as cost, time, and skill level required. States have expressed the need for a quick, easy, and inexpensive way to assess methane emissions from an orphaned well. The use of flux chambers can provide highly accurate data, but this approach can be time consuming and complicated to implement. Commercial high flow samplers may be more expensive than desired. There are many commercially available methane concentration detectors that vary in cost, but a concentration measurement does not provide an emission rate. To address these issues, the WP1 team developed and tested a methodology for quick and inexpensive methane emissions estimates from orphaned wells that utilizes downwind concentration measurements and wind field measurements, coupled with a plume dispersion model, to estimate methane emissions. Although the methodology does not provide the same high level of accuracy of flux chambers or high flow sampling, it allows for easy identification and classification of wells that are not leaking, leaking at low or mid-rates, or high emitters that may warrant additional investigation. With this methodology, wells that should be high priority for plugging due to high methane emissions can be quickly and inexpensively identified.

7.3.2 Best practices: Procedure for locating oil and gas wells in the Appalachian Basin that utilizes a compilation of publicly available digital resources

The top data gap or need as reported by the states through the IOGCC survey is orphaned well locations. Many wells may have been drilled and abandoned before accurate records were kept of their locations. Even when records of well locations are available, the accuracy of the geographical coordinates may not meet today's GPS standards and errors in location can be more than 50 meters, which makes finding the site in the field very difficult (Saint-Vincent et al., 2020). Surface indicators of a well's location (e.g., derricks, wellheads, tanks, and flow lines) are likely to have been removed and any remaining markers denoting the location of abandoned wells are easily overlooked in areas of heavy vegetation.

Well location efforts generally start with an investigation of recorded information such as well plats, county, and company development maps, as well as state agency well maps and databases. Due to the potential for these records to be incomplete and/or inaccurate, other publicly available resources can be used for identifying potential well sites, such as historic aerial photographs, topographical maps maintained by the United States Geologic Survey (USGS), state Light Detection and Ranging (LiDAR) surveys, and other historical maps.

The CATALOG team has reported a methodology for systematically compiling publicly available digital data at progressively smaller scales prior to embarking on field campaigns. Geographical Information System (GIS) technology is used to analyze digital data and create maps to guide field work activities based on the goals of a specific project. This workflow was developed in the Appalachian region and although certain aspects may be unique, the general process should be applicable to locating undocumented wells in other regions. The report of the methodology was designed to assist states, particularly in the Appalachian Basin, in their initial investigations of well locations for an identified area of interest.

7.3.3 Best practices: Aerial magnetic surveys to find well sites

Over the past decade, the U.S. Department of Energy's (DOE) National Energy Technology Laboratory's (NETL) has shown that airborne magnetic field surveys are able to detect and locate oil and gas wells, including undocumented orphaned wells. Because the infrastructure and well construction are primarily made of metal, the relatively large magnetic field they produce can be measured using a magnetometer. Well casing and other infrastructure containing high amounts of ferromagnetic alloys will then appear as strong magnetic field anomalies compared against the Earth's inducing field. This allows for the measurement of the magnetic field of infrastructure from low flying aircraft (approx. 100 ft above ground level (AGL)). The use of aerial magnetic surveying provides accurate locations for metallic-cased wells, which is of particular interest in areas where documented well locations are expected to be inaccurate or there is potential for the presence of undocumented wells. The NETL team completed a report of best practices for conducting aerial magnetic surveys for well finding. The sharing of the report will allow stakeholders to benefit from the decade of experience that NETL has in testing and demonstrating this technique.

7.3.4 Best practices: Selection of field testing and demonstration sites

Some research activities conducted under work packages in the CATALOG program will require the field testing and demonstration of various technologies and methodologies to verify their performance. To ensure that field testing and demonstration can be effectively conducted, the WP5 team has focused on: (1) assigning coverage of sites across the U.S. such that technologies and methodologies are relevant and successful in a variety of settings (different geology, topography, land use and cover, etc.), and (2) matching the needs of specific technology field tests and demonstrations with an available site.

7.3.5 Best practices: Defining key terms for the CATALOG project

Various terminology may be used to describe the disposition of non-producing oil and gas wells depending on state. Use of differing terms and definitions can hinder communication understanding of state and project needs. The WP5 team identified a need to address this communication problem by investigating terminology and definitions for each state. WP5 created a glossary ("Key Terminology for the CATALOG Project") summarizing key, relevant terms and associated generalized definitions – noting certain synonymous terms parsed out by state – to help streamline discussions and ensure consistent understanding across the project.

7.4 Data Summary and Accomplishments

7.4.1 Best practices: Use of downwind methane concentrations measurements and wind field measurements to model and estimate orphaned well methane emissions

Methodology and data sources are more completely described in Dubey et al. (2023). In summary, a Gaussian plume dispersion model was developed using a variety of atmospheric stability classes and wind speeds to estimate the expected methane concentration up to 10m downwind of a well emitting methane at a rate of 1 g/hr. The model was then tested via controlled releases and at several wells in New Mexico where methane emission rates were known and could be compared to modeled estimates.

Accomplishments

Figure 7.1 shows modeled methane emission rate as a function of wind speed and methane concentration measured 2m downwind of the source under stable atmospheric conditions. The graph illustrates how this approach can be used by operators to roughly estimate methane emission rate from an orphaned well by simple collection of wind speed data and downwind methane concentration, which is less time consuming and expensive than some alternative approaches, such as high flow sampling and flux chambers. The methodology is directly responsive to stakeholders' expressed need to measure methane emissions from orphaned wells in a way that is quick, easy, and inexpensive.

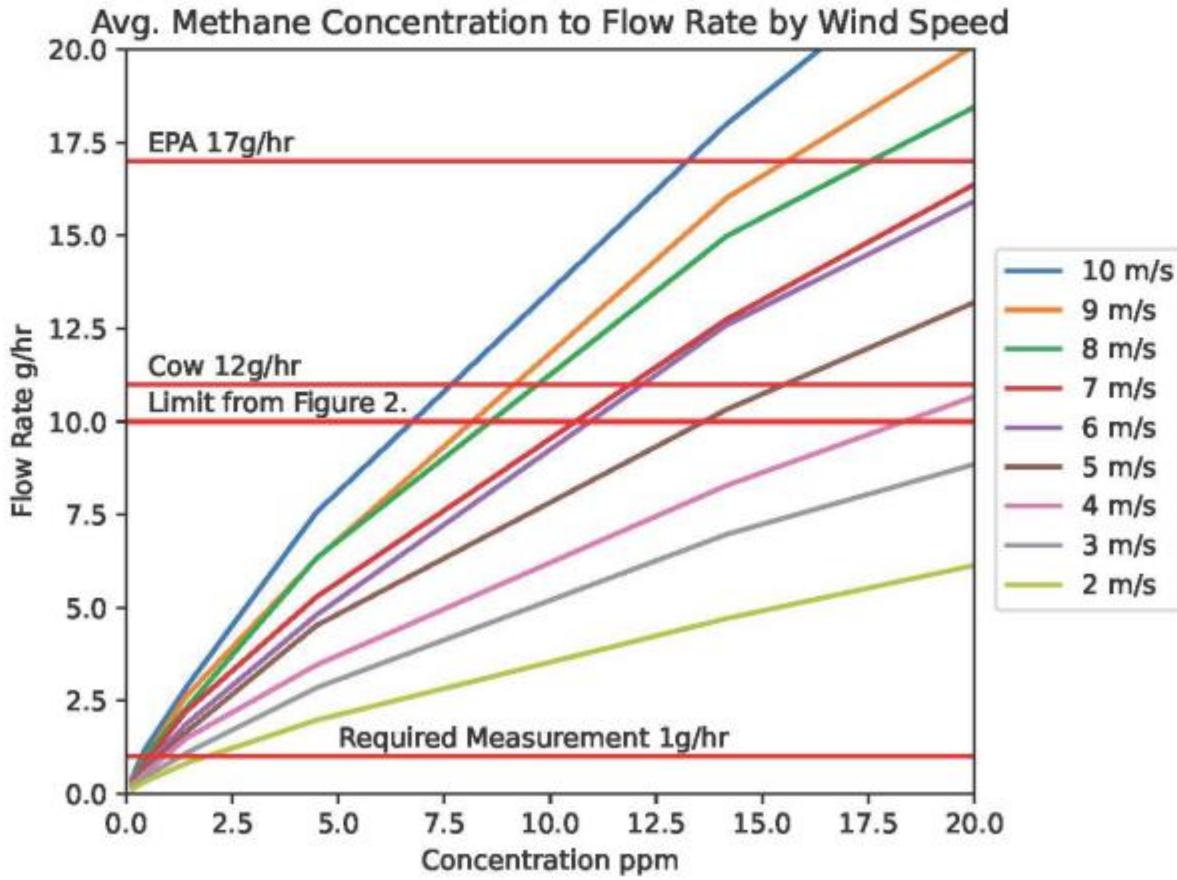


Figure 7.1: From Dubey et al. (2023). Methane emission flow rates are shown as a function of measured concentration 2m from the source and wind speed. Horizontal red lines are supplied for context; 1 g/hr is the methane reporting requirement for the BIL 2021, EPA 17 g/hr is the limit of detection using optical gas imaging to detect methane leaks.

7.4.2 Best practices: Procedure for locating oil and gas wells in the Appalachian Basin that utilizes a compilation of publicly available digital resources

Data sources and their links for the well locating methodology are listed in the report, "Procedure for locating oil and gas wells in the Appalachian Basin," which will be posted on the CATALOG website (Reeder et al., 2023). They include well location/production databases for each state in the Appalachian Basin (PA, OH, NY, KY, WV, TN, and AB), national databases curated by the USGS as well as three privately owned databases, USGS topographical maps, United States Department of Agriculture (USDA) aerial photographs, and historical mine maps from Office of Surface Mining Reclamation and Enforcement. Links for each state for LiDAR and other GIS data are also listed. The methodology for compiling the digital data involves using GIS software to develop a map application that layers well site location information from each of the data sources.

Accomplishments

The approach described in Reeder et al. (2023) has been developed and tested by the NETL team over the past 5+ years to locate potential well sites in areas that include Hillman State Park, PA; Oil Creek State

Park, PA; Daniel Boone National Forest, KY; an area around Oolagah Lake, OK; Wayne National Forest, OH; Stonewall Jackson Lake Resort Park, WV; and a private property in Olean, NY. Figure 7.2 shows as an example the layering of multiple digital data resources into a single map for the Olean, NY property that allows comparison of locations and provides a guide for field-based well location verification. This approach can be used as a first step in identifying potential well sites in an area of interest to improve efficiency in conducting ground-based field verification of well locations. Making the methodology publicly available through the CATALOG website for stakeholders' access allows them to adopt it for their own well finding purposes, providing them with the benefit of years of experience in using digital resources to facilitate well finding activities.

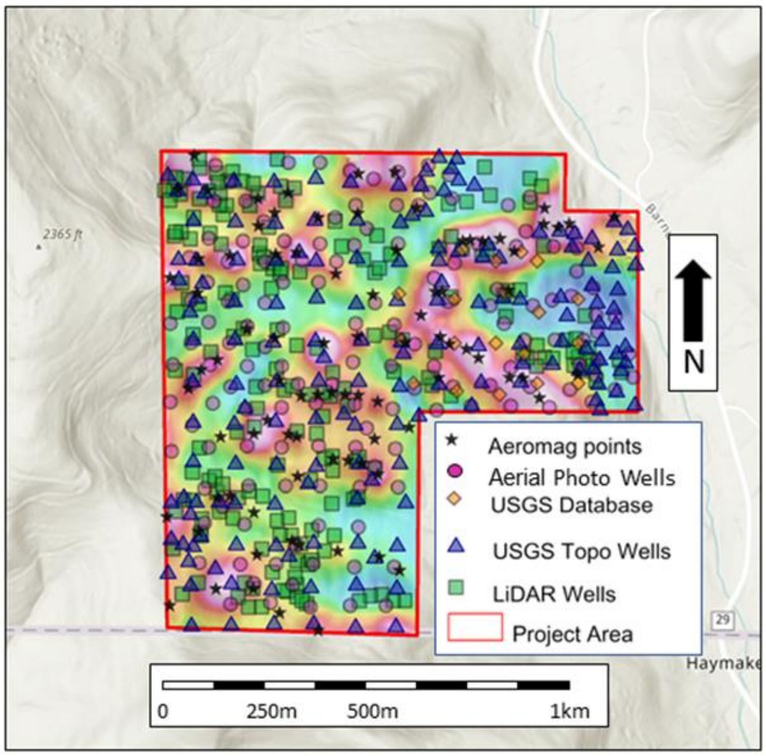


Figure 7.2: From Reeder et al. (2023). Data compilation at a project site in New York State.

7.4.3 Best practices: Aerial magnetic surveys to find well sites

Aerial survey design recommendations are based on empirical testing through field campaigns conducted at several locations over more than a decade: Hillman State Park, PA; Oil Creek State Park, PA; Ole Bull State Park, PA; MCC Partners Well Pad, PA; Salt Creek Oil Field, WY; Teapot Dome Oil Field, WY. Data processing, analysis, and filtering recommendations are also based on these same field campaign experiences as well as key literature resources (described more fully in Hammack et al., 2023). Sources of error in the magnetic data are identified as resulting from proximity of the sensor to metallic objects (i.e., the aircraft), temporal variation of Earth's magnetic background, and other time-varying effects such as solar storms, powerlines, and currents in the aerial platform that can interfere with the magnetic background.

Accomplishments

Best practices for aerial magnetic surveying to find potential well sites based on over a decade of experience from the NETL team have been summarized in Hammack et al. (2023) (Table 7.2). This technology is incredibly useful in finding metallic-cased wells, particularly documented wells with inaccurate/imprecise coordinates and undocumented wells. However, experience and science-based justification for survey planning and execution are necessary for successful application of the technology. Hammack et al. (2023) provides guidance for survey planning, flight parameters, data processing, and data analysis to guide others' efforts and maximize the potential for successful execution. Figure 7.3 shows an example of magnetic data collected via aerial platform at different aircraft heights and survey grid spacing. Tighter grid spacing and lower elevations give the best resolution of the magnetic data but have higher cost. From empirical testing, NETL has determined that a maximum flight height of 45 m and an inline spacing of 40 m can be used without significant degradation to the detection success rate. Above these numbers, the magnetic field loses the high frequency content and resolution necessary to reliably interpret the magnetic field from well casings.

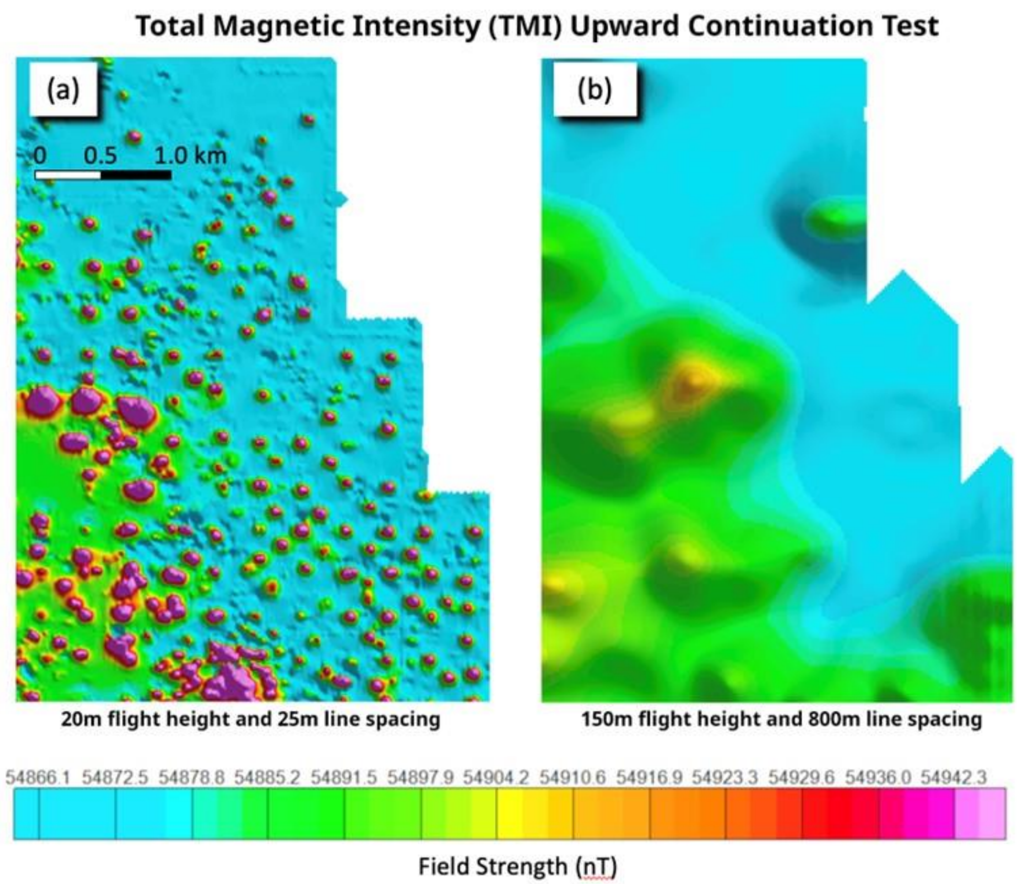


Figure 7.3: Excerpt of results from an aerial magnetic survey at Teapot Dome Oilfield, WY (from Hammack et al., 2023). Each map shows magnetic intensity at different aircraft heights and inline grid spacing during the survey. Magnetic anomalies that indicate well locations lose resolution as height and grid spacing increase.

7.4.4 Best practices: Selection of field testing and demonstration sites

To address coverage of field testing and demonstration sites across the U.S., the CATALOG team investigated the states with the highest populations of orphaned wells (Figure 7.4) and suspected undocumented orphaned wells and then matched states/regions with labs that either had a proximal advantage or had prior experience in working with the state agency or state oil and gas data. The assigned states/regions would then be the focus of that laboratory's field testing and demonstration activities. The regional/state focus for each of the national laboratories is as follows: NETL–Pennsylvania, New York, Ohio, West Virginia, Kentucky (Appalachian Basin); Lawrence Berkeley National Laboratory (LBNL)–California, Alaska, Texas (Permian Basin), Oklahoma, Colorado; Sandia National Laboratory–Oklahoma, Texas, New Mexico, California; Lawrence Livermore National Laboratory (LLNL)–California, Pennsylvania; Los Alamos National Laboratory (LANL)–New Mexico, Colorado (San Juan Basin), Utah, Wyoming. Except for Kansas and Louisiana, at least one national lab will be testing and demonstrating technologies and methodologies within the states with the ten highest populations of documented orphaned wells. While not meant to exclude a national lab from communicating with or conducting research in states outside their area of focus, this approach gives states resources to directly communicate and work with a specific national laboratory such that their unique needs are being addressed.

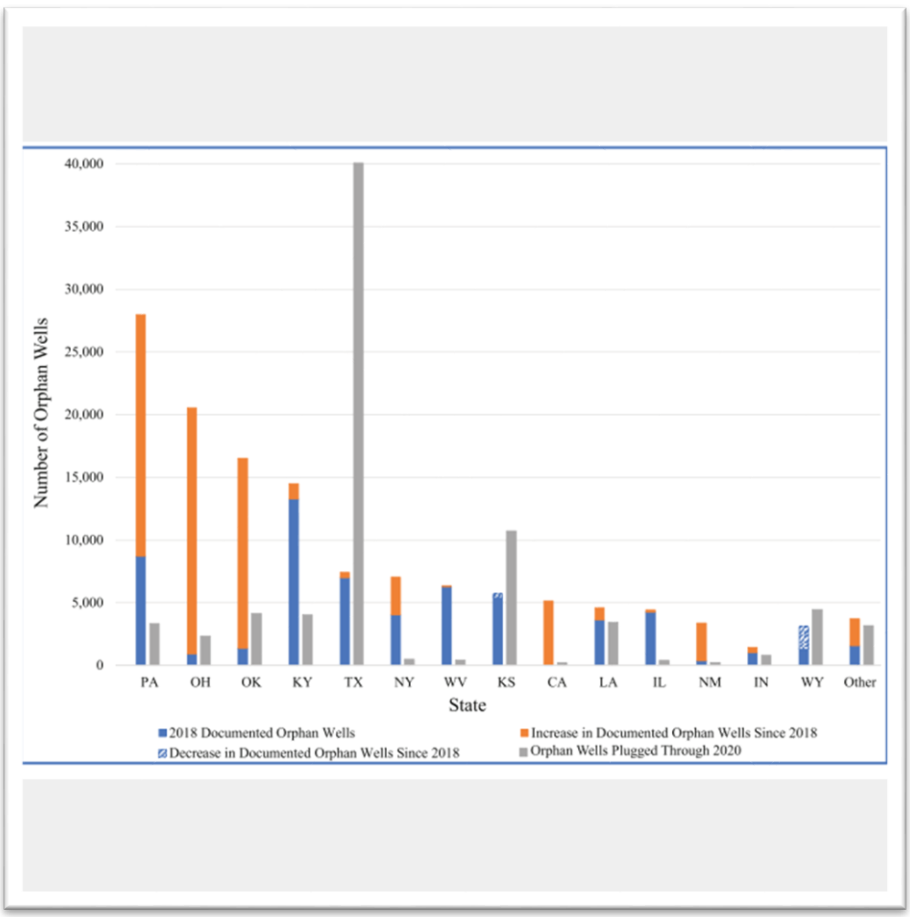


Figure 7.4: Documented orphaned wells as of IIJA enactment, from IOGCC (2021).

Figure 7.5: Field testing and demonstration focus areas for the five national labs participating in the CATALOG program.

To match the needs of specific technology field tests and demonstrations with an available site, the WP5 team developed a questionnaire to distribute to principle investigators (PIs) of distinct CATALOG research projects that would require field testing and/or demonstration. The questionnaire was divided into three primary sections including general overview, specific characteristics and final reflections. For each question, the PI was able to mark and explain the importance of each identified parameter that the WP5 team identified as potentially being important for a field test site. Aside from details about the research project, the following parameters were included on the questionnaire:

- Well configuration
    - Testing in a large region with many wells vs. testing at a single well
    - Prominence of the wellhead (e.g., buried, present, both)
    - Casing type (e.g., metallic, non-metallic)
- Well characteristics
    - Leaking or not leaking methane
    - Produced water storage; salts from produced water disposal/spills around well
- Geology
    - Depth of well or specific formations
    - Type of well (oil, natural gas, coalbed methane)
    - Density of wells in the region
    - Active wells in the vicinity (active field)
    - Presence of Naturally Occurring Radioactive Material (NORM)
    - Soil type/moisture content
- Geography
    - U.S. region/basin
    - Land ownership (private, public, tribal land, etc.)
    - Proximity to a national lab
    - Climate or season
    - Terrain (mountainous, hilly, flat, steepness of slopes)
    - Vegetation type (forested, little to no vegetation), suburban or urban, etc.

- Proximity to a disadvantaged community
- Proximity of water bodies (presence or absence)
- Proximity to sensitive ecosystems
- Proximity to occupied dwellings
- Ability to fly a drone (no regional prohibitions)

Accomplishments

Hillman State, Park, PA has been identified as a field testing and demonstration site that meets the needs indicated by several of the PIs for various CATALOG research projects (Table 7.2). The park provides 3,600 acres of public lands with approximately 190 well sites, most of which are undocumented. NETL and partners have conducted aerial magnetic surveys at the park starting in 2014 (manned helicopter and drone), and methane emission rates have been measured at 31 wells in the park using high flow sampling, optical gas imaging, flux chambers, and continuous monitoring. There are a variety of well types (buried, open hole, open casing cut off at grade, intact wellhead, etc.) and all the wells are currently unplugged. The park is suitable for ground-based or terrain-draped drone geophysical surveys as well as for training of stakeholders in well finding and characterization approaches. Additionally, the WellDone Foundation (https://welldonefoundation.org/) has been identified as a collaborator in providing well sites for emissions measurement collection and testing. They have had and will continue to have field sites across the country.

7.4.5 Best practices: Defining key terms for the CATALOG project

Multiple peer-reviewed journal articles, shared project resources, and governmental reports were consulted to assemble the glossary. These include: Boutot, J. et al. (2022), IOGCC (2021), U.S. Department of Interior (2022), and U.S. Environmental Protection Agency (2022).

Accomplishments

A report of differences in state-to-state definitions and development of a common understanding of terminology was completed, with the intention of guiding the national discussion and allowing direct comparison of research conducted in multiple states (Hora and Gunda, 2023). The key terms defined include abandoned, idle, orphaned (documented and undocumented), and a new term was introduced, "uncharacterized orphaned well," which describes an orphaned well whose location is known, but other properties or characteristics are unknown (Table 7.1). A quick-guide glossary of important terms used to categorize non-producing oil and gas wells facilitates common understanding and aids the project in engaging with state agencies that may use different terms. Additionally, efforts are underway to explore how the glossary could help inform the crosswalk for WP3 database development.

Table 7.1: Summary of Key Terms associated with CATALOG Project. Key synonyms from Boutot et al (2022) are included, especially for states being considered for project field testing activities.

Term	Key nuances	Notes
Abandoned	Non-producing well that is either plugged or unplugged.	**Regional Synonyms:** Deserted, dormant, idle, inactive, long-term idle, shut-in, suspended, temporarily abandoned, orphaned

Orphaned Well	An unplugged abandoned well with an unknown/insolvent operator such that a governmental agency is responsible for decommissioning and remediation.	**Regional Synonyms:** Abandoned, forfeited, revoked, shut-in, unknown CA: deserted, potential orphan NY: unknown, unknown located, unknown not located PA: DEP orphan, DEP abandoned UT: orphan-no responsible operator
Documented Orphaned Well	An orphaned well whose details (location and characterization) can be found in federal inventories or in a state database after undergoing a verification process, which can vary by state.	
Undocumented Orphaned Well	An orphaned well that has an unknown or unverified location and lacks formal documentation (e.g., drilling, completion, or inspection reports) establishing the existence and properties of the well.	**Regional Synonyms:** None; Lost; Unknown
Uncharacterized Orphaned Well	An orphaned well whose location is known, but well has unknown properties or characteristics.	Term introduced by CATALOG project

Table 7.2: Milestones for WP5.

Milestone List						
#	**Name**	**Description**	**Lead(s)**	**Due Date**	**Completion Date**	**Status**
5.1.1	Data Integration	Develop initial framework for well data integration	Pekney Enstrom	2/28/23	N/A	This activity has been moved to WP6 and WP8.
5.2.1	Field Demonstration	Establish stakeholders that will assist with field demonstration of integrated technologies	Pekney Enstrom	8/30/23	8/30/23	Multiple stakeholders and field demonstration sites have been identified (Hillman State Park, PA; Well Done Foundation; several state O&G agencies)
5.3.1	Best Practices	Develop outline and basic framework for best practices for well finding, emissions	Pekney Enstrom	8/30/23	8/30/23	An outline for best practices guidance was completed. Several best

		measurements and well integrity testing				practices documents have been completed and will be posted to the CATALOG website.

7.5 Future work

WP5 activities projected to continue/start in year 2 of the CATALOG program are listed below.

- Disseminating research results into best practices guidance
    - To make research products consumable on multiple levels, the WP5 team will communicate with WP leads, collect information and results, and share them in a format that is appropriate for our stakeholders on the CATALOG website. In year 2 expected website postings include:
        - Use of downwind methane concentrations measurements and wind field measurements to model and estimate orphaned well methane emissions: The methodology will continue to be tested, and findings and guidance on the approach will be shared.
        - Procedure for locating oil and gas wells in the Appalachian Basin that utilizes a compilation of publicly available digital resources: The methodology will continue to be tested and expanded for other regions of the U.S.
        - Best practices for aerial magnetic surveying for well finding: recommendations will be summarized from the journal article (Hammack et al., 2023)
- Automation of the procedure for locating oil and gas wells that utilizes a compilation of publicly available digital resources: There are opportunities to automate this process using machine learning techniques. WP3 has initiated efforts to pick well features from topographical maps using ML. Results will be shared via journal articles (in development) and on the CATALOG website.
- A procedure for assigning a probability or risk factor to potential well sites will be developed as a follow-up to the process for compiling digital resources to find well locations. Based on agreement between various resources, potential well sites will be assigned a higher or lower confidence in the accuracy of the location.
- Identification of field testing and demonstration sites will continue with ongoing communication with project PIs.
- Field testing and demonstration of deploying drones with a multi-sensor payload, started late in year 1, will continue in year 2. Ultimately the results of the testing will be shared in best practices guidance in a future year.

8 Work Package 6- Data Management

8.1 Section Summary

The RDD&D efforts performed under this Program will generate large volumes of data and information that are integral to the research, knowledge development, collaboration, and, ultimately, the successful identification, characterization, and plugging and abandoning activities by State, Federal and Tribal agencies. As a result, the effective collection, storage, and preservation of these data are critical in enabling successful collaborations, the development and testing of novel technologies, and reliable access to data, products, and information created from this Program. The overall goal of WP 6 is to provide access to information and data for the Program and support the data lifecycle and management for the Program to ensure enduring, secure access and sharing of data across all work package (WP) activities and with collaborators, including State, Federal and Tribal agencies.

8.2 Activity

Work in Year 1 for WP 6 broke down into three main categories: 1) development of a program-based interactive online application, 2) documentation of data standards and quality requirements, and 3) the integration and storage of data, datasets, and research products in EDX. A summary of key activities for Year 1 are listed below for each of these three categories:

- 1) Development of a Program-based interactive online application
  - Developed requirements for and create an online web application that provides members of the research consortium and the public with details of the Program, on-going WP activities, related activities and resources, and any research products.
  - The initial website was made live in Jan. 2023, to support sharing information on the Program, as well as to aid in collecting and building out content for other areas of the website, including information on the WP focus areas and resources.
  - Requests to simplify the URL were perdues, and the website was rolled over to the current URL for public access in May 2023.
  - Since the website's release in Jan. 2023, information on the Program and seven focus areas have been added. Additionally, five Program resources have been added to the website (1 publication, 2 presentations, and 2 datasets).
  - Since the release of the CATALOG website at https://catalog.energy.gov/ in May of 2023, the website has been visited more than 3,000 times by over 1,300 different users from across the world (Figure 8.1).
    - Majority of the users are accessing the website directly and are from the United States, but some are discovering the website through internet search and website recommendations.
    - Most visits have focused on the main page about the CATALOG program, but the Focus Areas pages and some of the resources have also been some of the top visited areas of the website.
    - Analytics from the website show that more than 10,000 events, which are a measure of the interaction with the website, whether it's loading a page, clicking a link, or downloading a resource, have been captured since the websites release in May 2023.

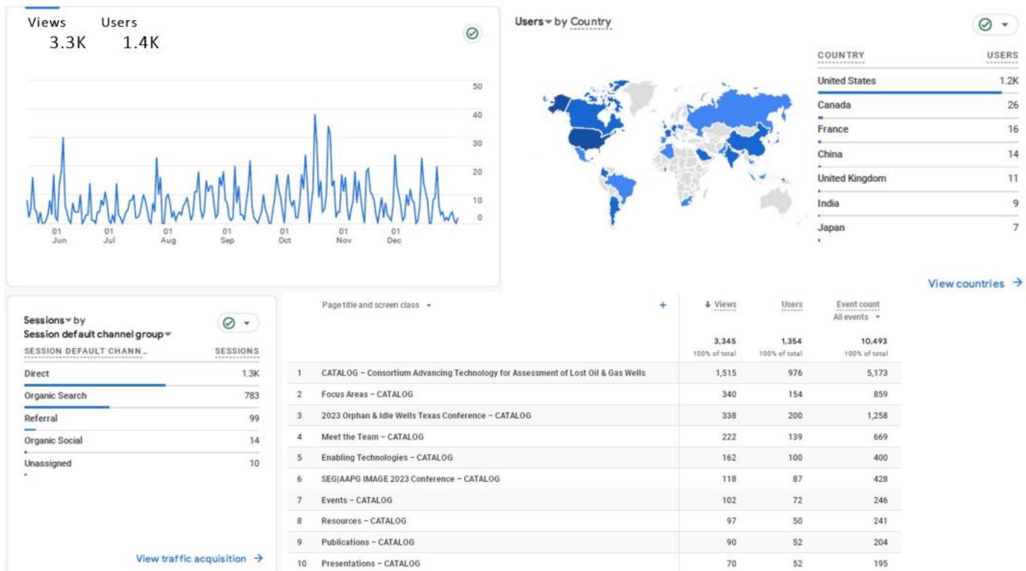


Figure 8.1: Summary statistics about users and viewers of the CATALOG website since it's public release in May 2023.

- 2) Documentation of data standards and quality requirements
    - The WP 6 team held various meetings throughout Year 1 to discuss and support Program activities related to anticipated data types and formats, discuss on any recommended or required standard(s) for use (e.g., file type, data structure, metadata fields, etc.) and any required information for each data standard needed to successfully document data quality that will be created or generated by each WP.
    - The WP 6 team also meet with multiple State and Federal agencies to learn about their data needs, discuss primary forms of access, and capture their data and activity needs or requirements. This information will be used to support decisions for sharing data from the Program in future years work.
- 3) Integration and storage of data, datasets, and research products in EDX
    - At the end of Year 1, almost 4,000 files and data products, totally ~100 GBs, have been uploaded to the Program's EDX Collaborative Workspace (restricted to viewing by team members only) for accessibility and discovery. These files have been downloaded more than 1,000 times for use by various Program members in support of their work.
    - While the Program works on Year 1 research, the WP 6 team began outlining documents and file structures on the Program's EDX Collaborative Workspace that will be available to their efforts in uploading, sharing and storing data, datasets, and research products on EDX, including reports, presentations, datasets, tools, code, and models. The WP 6 team will share these resources with Program members to help them unify data cataloging and improve in the overall discovery and appropriate usability of resources from the Program, as data and research products begin to migrate on the EDX collaborative workspace, as well as out for public access.

8.3 Motivation and Background

The motivation behind this WP is to facilitate data and information sharing and discovery along with provide enduring, secure access to data, datasets, and research products generated.

8.4 Data Summary

The online web application, referred to as CATALOG, was launched summer 2023. It provides members of the research consortium and the public with details of the Program, on-going WP activities, related activities and resources, and any research products from Year 1 activities.

8.5 Future Work

With the website establish, the WP6 team will continue to support updates to the CATALOG website, leveraging content provided and identified by the teams. This will include pulling research product citations and links from quarterly reports, as well as coordinate with the appropriate points of contacts from the program and each work package to ensure content on the website reflect current work, upcoming events, recent R&D products, and note program accomplishments.

Additionally, following the significant R&D progress from the year, majority of the WP teams are beginning to prepare to publish and share the results of their work. The WP6 team will be available to coordinate with different WP teams and provide additional support in preparing their data and data products for DOE review and public release through EDX, as appropriate. This may include working with the teams to support the publication of affiliated data and/or code required by publications, help identify data products, identify recommended data standards, document metadata, add links related public data and data products on EDX, group UOW related resources through EDX, identify appropriate reviewers and approvers, and discuss recommended data licenses and use restrictions. The WP6 team and EDX support team will be available to help identify potential solutions for WP teams from the program that are looking to generate or have created big datasets, those that required terabytes or petabytes of storage size, contain several hundreds of files, or require different upload/download solutions beyond the capabilities of common network connections. With data and research products available on EDX, within either the Program's Collaborative Workspace or released to the public, these products can be queried to generate lists of all data, reports, presentations, datasets, tools, code, and models generated from all WPs and uploaded to EDX throughout the entire lifespan of the Program.

9 Work Package 7- Data Extraction

9.1 Section Summary

The focus of Work Package 7 is data extraction from historic oil and gas regulatory records. This project began in August of 2023. The goal is to build an online platform that oil and gas regulators, industry operators, researchers, and other stakeholders can use to digitize information from scanned images of well completion reports and other important regulatory records (Figure 9.1). To date, the team has identified a work scope, begun conversations with potential tool users, acquired example documents from two states for tool development purposes, and tested various approaches for optical character recognition (OCR) and large language modeling (LLM). Work has shown that OCR and LLMs are powerful tools that have significant potential to accurately digitize information from well records. Development of the proposed tool has significant potential to reduce the effort required to build electronic databases of historic oil and gas information.

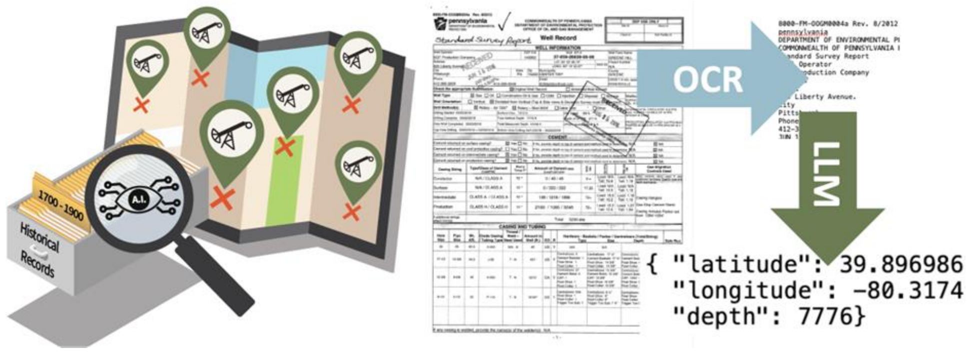


Figure 9.1: Information about inadequately documented oil & gas wells can be extracted from historical records using large language models. Here, we show information about the location and depth of a well extracted from a scanned document with complicating factors like stamps and handwritten notes. These forms sometimes have errors (such as the longitude having the wrong sign – indicating the wrong hemisphere of the globe) that large language models can automatically correct.

9.2 Activity

9.3 Motivation and Background

An understanding of the design and construction of a well is critical to ensure effective plugging and to accurately quantify the benefits of remediation. Well design and the materials in place may also impact the success of remote well detection technologies. State oil and gas regulatory agencies maintain historic regulatory records for known wells under their jurisdiction. These records contain valuable information about the design, construction, and operation of wells over their lifetime. Aggregation and analysis of this information for known wells could provide the only opportunity for insight into the characteristics of undocumented wells in their vicinity. While most state oil and gas agencies have digitally scanned historic well records, many lack the funding and bandwidth to manually extract information from them. Consequently, public (and private) databases of oil and gas information generally lack well construction details. Computational approaches that employ OCR have potential to rapidly extract data from older scanned documents. However, older records are often difficult to interpret due to faded text, stamps obscuring essential data, or handwritten entries. Combining OCR with machine learning approaches such as LLMs is potentially the key to overcoming these challenges. We are developing an online tool that leverages both OCR and LLMs to digitize information from oil and gas records. Our goal is to provide technology to stakeholders in the orphaned well community to help facilitate the digitization of information from historic documents.

Activities under Work Package 7 include:

- 7.1. Stakeholder engagement, data collection, and document sorting/labeling.
- 7.2. OCR and LLM development
- 7.3. Tool development

All laboratories involved in the CATALOG project are engaged in the activities of Work Package 7. Responsibilities for each lab are shown in Table 9.1.

Activity #	LANL	LBNL	LLNL	NETL	SNL
7.1					
7.2					
7.3					

Table 9.1: Activity Leads (Blue) and Activity Partners (Tan*)*

9.4 Data Summary

Data collection efforts under Work Package 7 focus primarily on oil and gas well records. Both scanned and electronic well records will be valuable for the effort. Scanned images are necessary to develop effective approaches for OCR. It is easier to accurately extract information from electronic records. These data will be valuable for validating OCR methods and training LLMs. Work Package 7 researchers have collected all (~1.5 TB) well records from the state of California and all (~10 GB) well completion reports from Colorado for the project. Small batches (~1K) of well completion reports have also been gathered from Pennsylvania and West Virginia. Additionally, digital data describing the construction of oil and gas wells in Colorado was acquired by Work Package 7 researchers and will be used to validate OCR and LLM tools applied to Colorado records.

The Work Package 7 team applied three widely used OCR tools – Tesseract, Easy OCR, and Google Document AI – to 162 drilling completion reports of varying quality. Google Document AI was found to be the most accurate, extracting depth information correctly from 154 (95.1%) of the 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%) of these records. The 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.

The Work Package 7 team has also achieved success applying LLMs to relatively clean well records using a two-step approach: first, applying OCR to convert the form into text, and then interpreting the text with LLMs. For example, the team accurately identified the latitude and longitude with 100% accuracy from 150 drilling completion reports containing location, depth, time, and other information in PDF files submitted to the Colorado Energy and Carbon Management Commission. The team applied a Zero-shot learning approach on two models, DocQuery and ChatGPT 3.5. Both models were able to rapidly extract well locations from 150 well completion reports from the state of Colorado. DocQuery, an older model, struggled on a more complex and smaller set of 10 reports from Pennsylvania (see Fig. 2). However, ChatGPT was able to extract the latitude, longitude, and depth from these forms accurately. This test demonstrates the unique capability and advantage of modern LLMs for pulling structured information out of semi-structured data. It remains an open question as to how complex and old the forms can be before this approach breaks down.

9.5 Future Work

Historical records are valuable resources in the quest to address the environmental challenges posed by orphaned wells. By harnessing OCR and advanced machine learning techniques, states and regulators can access crucial data to inform their well identification and management strategies. Using such automated computational approaches costs significantly less than field-based approaches and can be scaled more readily across the United States. Thus, the potential exists to use the tools developed under Work Package 7 to greatly reduce the cost of characterizing under-documented orphan wells.

Future efforts under this Work Package will focus on three areas: 1) stakeholder engagement and tool development, 2) data collection and sorting, and 3) OCR and LLM development. Multiple state regulatory agencies have expressed interest in the tool under development in Work Package 7. We are planning a series of meetings with these stakeholders to identify their needs for historic document digitization to help guide the development of the tool. After these conversations, designs of the tool will be generated and iterated on with stakeholders. Once a design approach is selected, a prototype will be created by the end of year one for the Work Package. Data collection efforts will continue to focus on gathering well completion reports from states. We hope to gain direct access to records through relationships with interested state agencies. However, web scraping and document mining approaches are also being developed in parallel to ensure that the necessary data are available to the team. Efforts to develop machine learning techniques for document identification and sorting are also under way and will be matured over the next year. With preliminary testing of OCR and LLMs complete, the team will identify programs/algorithms to use for tool development. For OCR, templates for data extraction will be developed and the methodology will be refined. LLMs will continue to be trained with more data as they are collected.

10 Work Package 8- Composite Database and Dashboard

10.1 Section Summary

The focus of Work Package (WP) 8 is to build a nationwide composite database that integrates disparate data and provides information that can be used to identify orphaned wells, UOWs, and wells that were previously undocumented but have since been documented through field or data validation methods. Specifically, undocumented orphan wells that are identified through methods such as historical record mining, magnetic surveys, etc., will be added to the database once confirmed by in-person well confirmation checks.

10.2 Activity

WP8 builds off and expands upon prior work done in support of WP3, which worked to develop an integrated dataset consisting of private and public well data sets and select attributes that can be used to assist in identification of UOWs. In year 1, the team built off work from the CO2-Locate database (Romeo *et al*. 2023), which has developed a baseline method and database schema for integrating geospatial well data from multiple, authoritative sources. Working with this baseline dataset, the team worked to further integrate additional public well data resources from state and tribal regulatory authorities, as well as review database attributes and ensure consistency with terms and standards being developed by WP5. The NETL team also began work to review state regulatory databases to identify any records that represent now verified UOWs, and started outlining how these records can be documented within the database to meet the goals and objectives of this WP.

10.3 Motivation and Background

To support the development of tools, technologies, and guidelines for identifying UOWs, it is often necessary to begin by understanding where the known wells are for a given area. Most of the current oil and gas well data and information are managed at the state level, and each state has different data reporting requirements and methods to manage and maintain their system of records. Since oil and gas reservoirs often extend across state boundaries it is necessary to build composite interstate datasets for proper basin or field investigations. Although this is a common practice in industry, there is no such composite database in the public domain, with explicit spatial locations, that is currently available and yet such a database is often needed to support on-going analytics, models, tools, and field-based data collection surveys that can help identify UOWs.

10.4 Data Summary

Currently, the WP team has integrated more than 50 publicly available data resources, which provide more than 4 million well records. The team has also processed national well data from proprietary databases to provide derivative, spatial summaries for more than 4.6 million wells across the U.S., providing summaries related to wellbore spud dates, depth, status, etc. As of now, a subset of these data is available for public use and download through version 1 of CO2-Locate (Romeo et al., 2023; https://edx.netl.doe.gov/dataset/co2-locate). Version 1 database provides more than 2.1 million well records from public data sources, from Alaska, Alabama, Florida, New York, Pennsylvania, South Dakota, Utah, Virginia, and West Virginia, plus some national data from the Homeland Infrastructure Foundation-Level Data (HIFLD) (Figure 10.1). The WP 8 team has since added ~2.5 million well data records. This includes the integration of databases from state regulators that relate to orphaned wells, including the USGS' Documented Unplugged Orphaned Oil and Gas Well Dataset (https://www.usgs.gov/data/united-states-documented-unplugged-orphaned-oil-and-gas-well-dataset).

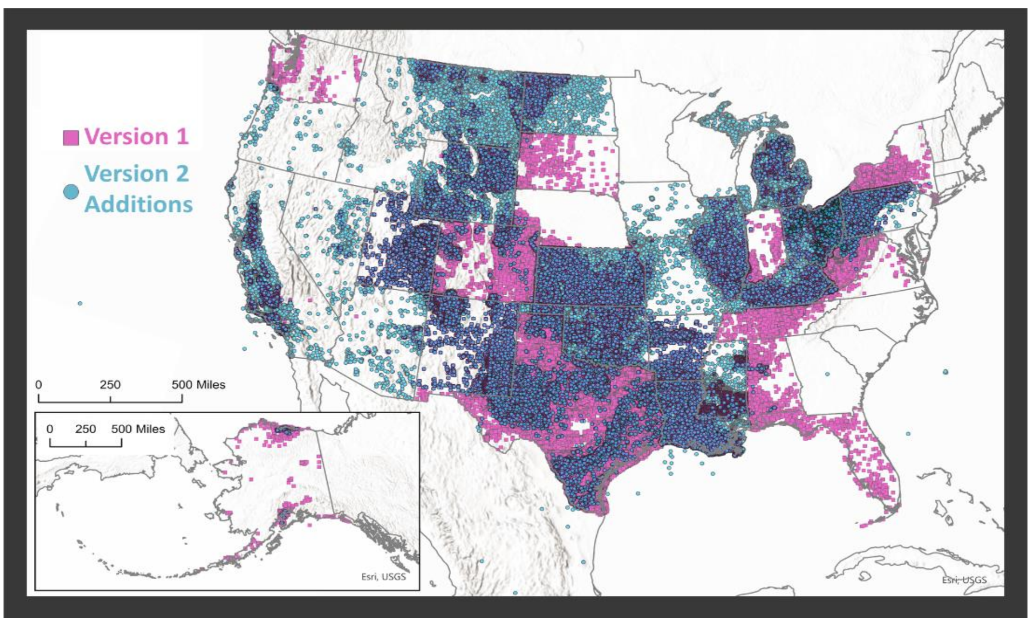


Figure 10.1: Comparison of wellbore records obtained for version 1 of the CO2-Locate database (available for download on EDX, https://edx.netl.doe.gov/dataset/co2-locate) and additional wellbore resources added for version 2 of the database, with a public release anticipated in Spring 2024.

10.5 Future Work

The team will continue working on data collection, processing and integration to support future updates and release of wellbore data to assist with UOW efforts. The team has also begun designing and developing a web-based mapping application to provide access to and visualization of the composite well database, as well as support user-enabled filter, query, select, and download functions. An alpha version of this application is expected to be available for internal UOW WP teams to review and provide feedback on the application by the end of November 2023. Additionally, the team has begun working with other WP teams to identify related data collected or generated that can be added to the web mapping application as overlay layers (e.g., location of field verified wells, footprint of on-going field sampling and remote sensing efforts, predictive raster layer of undocumented wells, etc.). The NETL team will be working with these WP teams to continue to understand what future information or data might be produced from their field or analytics-based projects to better identify UOWs and use these conversations to continue to outline database integration schemas for capturing and presenting this information either directly within the integrated database or as associated data layers in the web mapping application for future work.

11 Work Package 9- Field Teams

11.1 Section Summary

WP9 Field Teams is an inter-lab group that will rapidly respond to UOW requests from federal, state, tribal, and local communities with the appropriate level of support. Emphasis is put on airborne sensing capabilities (electromagnetic, magnetic, methane concentration mapping) though ground-based surveys (methane concentrations and flow rates, well integrity) are also key support capabilities. WP9 will remain closely tied to Work Packages 1-3 and 5 since the field campaigns will deploy instrumentation and acquire data relevant to one or more of these WPs.

11.2 Activity

The Field Team has identified three large field areas that have high UOW potential: Four Corners region, Osage County, NM, and Stonewall Jackson State Park, WV (Figure 11.1). We are working closely with stakeholders (Navajo Nation, BLM, Osage Nation, West Virginia Division of Natural Resources) to ensure our field deployments meet all objectives and provide necessary support. These areas represent three distinct regions within the continental U.S. in terms of terrain, vegetations, and soil type. We hope to identify field areas in two additional regions to capture as broad a range as possible of environments in which UOWs are highly likely.

Our first field campaign is scheduled for the end of October 2023. Osage Nation and the Osage Mineral Council identified two ranches that are good candidates for airborne magnetic mapping (Figure 11.2). These areas are large enough to warrant multiple field deployments. The first trip will cover a square mile over one of these ranches. UAVs will be dual mounted with a magnetometer and methane sensor. The existing BIA O&G well database will serve as a reference for magnetic map interpretations. WP3 ML work on historical topographic map analysis has “located” O&G wells not in BIA’s database, which the Field Team will verify. A ground team will also work to verify ML results, perform ground magnetic surveys to support airborne data, and characterize methane point sources (leaky wells).

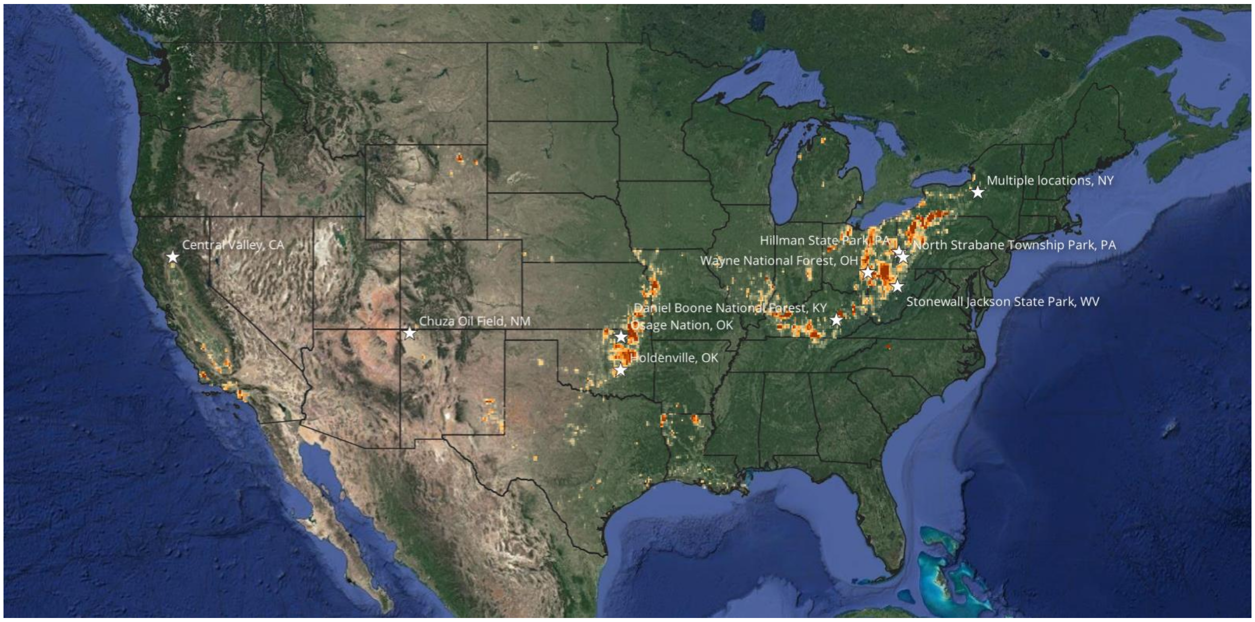


Figure 11.1: Areas of interest. WP9 is planning field campaigns at the Four Corners (Chula Oil Fields), NM, Osage Nation, OK, and Stonewall Jackson State Park, WV. Central Valley, CA is another potential area for WP9. Remaining stars have areas previously surveyed by NETL (pre-UOWP-CATALOG).

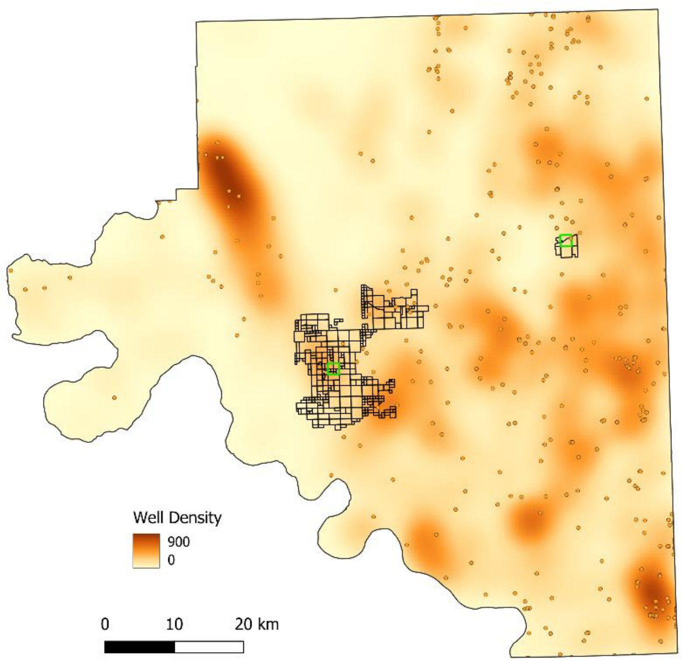


Figure 11.2: Osage County, OK. Well density (BIA O&G well database) and two priority properties (black polygons) identified by Osage Nation and the Osage Mineral Council. Focus areas (green polygons) on Osage Nation Ranch (west) and Lost Creek Ranch (east) are planned for the first of multiple field deployments. The orange dots are the results of WP3's Historic Topographic Map Classification algorithm for Osage County.

11.3 Motivation and Background

Following on OUWP-CATALOG's Best Practices and ML techniques, the program needs to test guidelines in the field and supply real world data sets to help train and test algorithms. Additionally, the programs should provide support to communities directly impacted by UOWs by demonstrating that the consortium can locate UOWs and characterize located OWs.

11.4 Data Summary

No field data has been acquired yet. Prior to field campaigns, during the planning phase, public geospatial data and existing well databases are being collected to physically characterize areas of interest and identify clusters of O&G wells where point source methane quantification may be efficiently done.

11.5 Future Work

Upon completion of the first field campaign in Oklahoma, WP9 will immediately begin data processing and packaging in tandem with planning the next field campaign. The next field campaign will either be a return trip to Osage Nation or the Four Corners.

Field work remains focused on drone-based magnetic anomaly and methane concentration mapping for the purposed of detecting undocumented wells. Additionally, each campaign will offer an opportunity for ground-based work—point source methane quantification, ground truthing ML results, and targeted well integrity characterization—that directly informs other work package efforts.

With each field campaign, the team will apply data standards defined by WP3 and provide feedback to the Best Practices Guidelines developed by WP5. Field data will be made available to CATALOG as a whole as soon as possible and published on CATALOG's website.

12 Appendix A - References

1. Beskardes GD, CJ Weiss, E Um M Wilt and K MacLennan, The effects of well damage and completion designs on geoelectrical responses in mature wellbore environments, Geophysics, **86**, E355-66 (2021).
2. Beskardes GD and CJ Weiss, Well integrity modeling with electric fields by using hierarchical geo-electric models, AGU Fall Meeting (2022).
3. Boutot, Jade, Adam S. Peltz, Renee McVay, and Mary Kang. 2022. "Documented Orphaned Oil and Gas Wells Across the United States." *Environmental Science & Technology*, September, acs.est.2c03268. DOI: 10.1021/acs.est.2c03268.
4. Downs, C., Eagleston, H., Oseghae, I. (2023, August 30). *Aiding the detection and identification of abandoned and orphaned oil and gas well with predictive modeling in the GIS environment* [conference presentation abstract]. International Meeting on Applied Geoscience & Energy, Houston, TX, United States.
5. Downs, C., Kadeethum, T., Heath, J. 2022, *Persistent homology-based feature detection from remote-sensing data*, Sandia National Laboratories, SAND2022- 12432R.
6. Dubey, M.; Meyer, A.; Dubey, M.; Pekney, N.; O'Malley, D.; Viswanathan, H.; Govert, A.; Biraud, S. (2023). How to estimate O&G well leak rates from near field concentration and wind observation. LA-UR-23-20659 Los Alamos National Laboratory https://catalog.energy.gov/wp-content/uploads/2023/05/20230103-OrphanWellLeakQuantificationFECM22.pdf García-Fernández, M., Álvarez-Narciandi, G., López, Y.Á. and Andrés, F.L.H., 2022. Improvements in GPR-SAR imaging focusing and detection capabilities of UAV-mounted GPR systems. ISPRS Journal of Photogrammetry and Remote Sensing, 189, pp.128-142.

7. Hammack, R.; Veloski, G.; Sams, J. Kohnke, C. (2023). Aeromagnetic surveys for the location of undocumented orphaned wells. Submitted to The Leading Edge, currently under revision. Summary available on request; will be posted on the CATALOG website.
8. Hora, P. and T. Gunda (2023). Key Terminology for the CATALOG Project. Available on request; will be posted on the CATALOG website.
9. Interstate Oil & Gas Compact Commission (IOGCC). (2021). "Idle and Orphan Oil and Gas Wells: State and Provincial Regulatory Strategies." Accessed online at: https://iogcc.ok.gov/sites/g/files/gmc836/f/iogcc_idle_and_orphan_wells_2021_final_web.pdf
10. Interstate Oil and Gas Compact Commission (IOGCC). 2023b. "Idle and Orphan Oil and Gas Wells: State and Provincial Regulatory Strategies. Supplemental Information on State Prioritization Systems for Orphan Wells." https://iogcc.ok.gov/sites/g/files/gmc836/f/prioritization_report_7.10.23.pdf.
11. Jol HM, editor. Ground penetrating radar theory and applications. Elsevier; 2008 Dec 8.
12. Kang, Mary, Jade Boutot, Renee C. McVay, Katherine A. Roberts, Scott Jasechko, Debra Perrone, Tao Wen, et al. 2023. "Environmental Risks and Opportunities of Orphaned Oil and Gas Wells in the United States." *Environmental Research Letters* 18 (7): 074012. https://doi.org/10.1088/1748-9326/acdae7
13. MacLennan K, G Nieuwenhuis, M Wilt and JM Pendleton, Evaluating well casing integrity with non-invasive electromagnetic methods, SPE Annual Technical Conference and Exhibition (2018).
14. O'Malley et al. (2023). The undocumented orphaned well challenge: An interdisciplinary opportunity to enhance sustainability.
15. Reeder, M. et al., (2023). Procedure for locating oil and gas wells in the Appalachian Basin. Submitted to The Leading Edge, currently under revision. Summary available on request; will be posted on the CATALOG website.
16. Romeo, L, Pfander, I., Sabbatino, M., Sharma, M., Amrine, D.C., Bauer, J., and K. Rose. CO2-Locate. 4/5/2023. https://edx.netl.doe.gov/dataset/co2-locate. DOI: 10.18141/1964068.
17. Saint-Vincent P., Sams J., Hammack R.W., Veloski G.A., Pekney N.J. (2020). Identifying abandoned well sites using database records and aeromagnetic surveys. Environ. Sci. Technol., 54, 13, 8300–8309. https://doi.org/10.1021/acs.est.0c00044.
18. Travassos, X.L., Avila, S.L. and Ida, N., 2020. Artificial neural networks and machine learning techniques applied to ground penetrating radar: A review. Applied Computing and Informatics, 17(2), pp.296-308
19. U.S. Department of Interior (2022). "FY 2022 State Initial Grant Guidance." Accessed online at: https://www.doi.gov/sites/doi.gov/files/state-initial-grant-guidance-bil.pdf
20. U.S. Environmental Protection Agency (2022). "Inventory of U.S. Greenhouse Gas Emissions and Sinks 1990-2020." Accessed online at: https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks
21. Weiss CJ, Finite-element analysis for model parameters distributed on a hierarchy of geometric simplices, Geophysics, **82**, E155-67 (2017).
22. Wilt M, Wellbore integrity mapping using well-casing electrodes and surface based electrical fields, Final Report for SBIR Phase I, DE-SC0015166 (2016).
23. Wu, K., Rodriguez, G.A., Zajc, M., Jacquemin, E., Clément, M., De Coster, A. and Lambot, S., 2019. A new drone-borne GPR for soil moisture mapping. Remote Sensing of Environment, 235, p.111456.