CATALOG – WP3 – Sensor fusion and data integration with Machine learning

LLNL FEW0285/FEW0299

Jacob Trueblood (lead)
Lawrence Livermore National Lab

Charu Varadharajan (co-lead)
Lawrence Berkeley National Lab

U.S. Department of Energy
National Energy Technology Laboratory
Resource Sustainability Project Review Meeting
April 2-4, 2024

Consortium Advancing Technology for Assessment of Lost Oil & Gas Wells.

catalog.energy.gov

LLNL-PRES-862255
This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Project Overview: CATALOG focus areas

CATALOG consists of 8 focus areas

Methane Detection and Quantification

Well Identification

Sensor Fusion and Data Integration

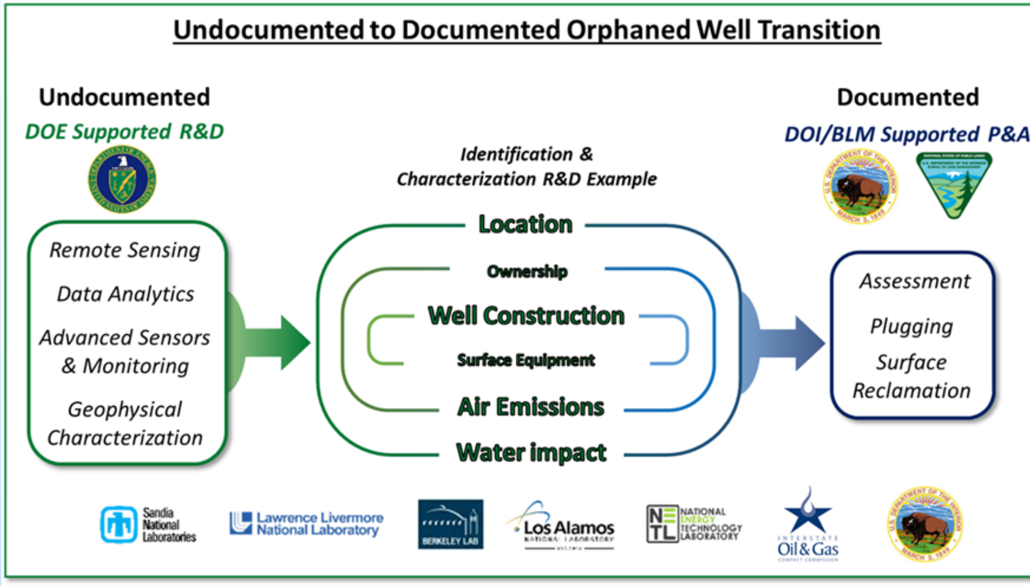
Well Characterization

Best Practices for States to Employ

Field teams

National Well composite database integration

CATALOG Data Extraction



This talk focuses on the sensor fusion and data integration focus area of the CATALOG program

Project Overview: Budget

- 5 Years $30M BIL funding, FY 2023, 2024 Appropriations $10M/yr
- October 1, 2023 – February 28, 2027

Los Alamos NATIONAL LABORATORY EST. 1943
BERKELEY LAB
Lawrence Livermore National Laboratory
NATIONAL ENERGY TECHNOLOGY LABORATORY
Sandia National Laboratories

Project Overview: Participants

Work Package 3 Participants

– Lead: Jacob Trueblood

– Co-Lead: Charu Varadharajan

LLNL	LANL	SNL	LBNL	NETL
Jacob Trueblood	Dan O'Malley	Christine Downs	Fabio Ciulla	Jennifer Bauer
Ken Enstrom	Hari Viswanathan	Holly Eagleston	Charu Varadharajan	Mumbi Mundia-Howe
Xianjin Yang	Mohamed Mehana	Teeratorn "Meen" Kadeethum		
Yuan Tian	Javier Santos			
Hewei Tang	Anastasiia Kim			
Jaisree Iyer	Ismot Jahan			
Claire Knight				
James Reimer				

Project Overview: Priorities

Work Package 3 Priorities for the year

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*

Project Overview: Activities

Work Package 3 Activities

1) *Gap analysis, developing requirements for hardware, data integration and ML (Priorities 1, 2, 3)*
2) *Hardware integration(Priority 1)*
3) *Evaluate data integration approaches (Priority 2 and 3)*
4) *Selection, training, deployment, and benchmarking (ML and other) of algorithms (Priority 2 and 3)*

Activity #	LANL	LBNL	LLNL	NETL	SNL
3.1					
3.2					
3.3					
3.4					

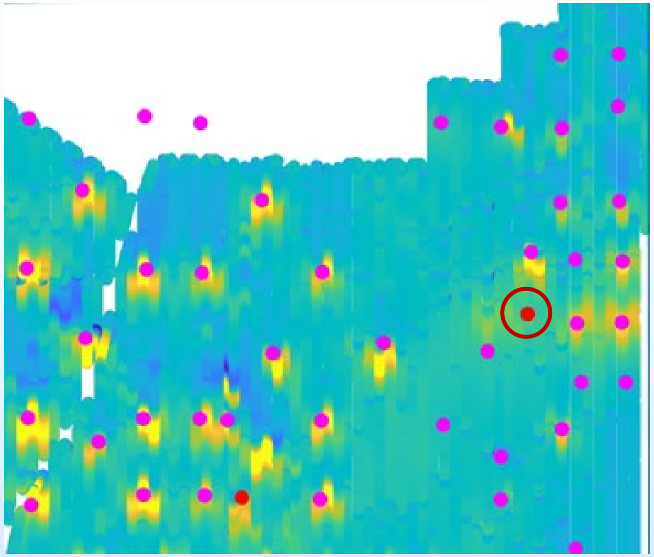
WP3 Activity Leads (Blue) and Activity Partners (Tan)

We are considering multi-sensor for integration for locating orphaned wells

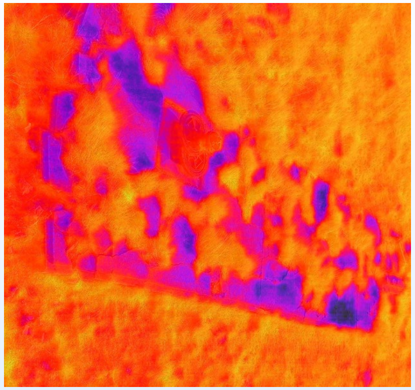
	Sensor	Artifact Detected	Area coverage
Manned Aircraft			
	Hyperspectral (7.4-11.8 µm)	Methane Plumes	Wide area search
	SAR (X-band)	Well heads/ infrastructure	Wide area search
sUAS platforms			
	Hyperspectral (350-1000 µm)	Well heads	Wide area search
	Lidar	Well heads	Wide area search
	RGB Camera	Well heads	Wide area search
	TDLAS	Methane Plumes	Localized search
	Magnetometer	Ferrous well casing	Wide area search
	Thermal	Well heads	Wide area search
	Methane Spectrometer	Methane Plumes	Localized search
	GPR (500MHz)	Well casing	Localized search

LLNL-PRES-862255

Goal to get to a robust set of sensor that can efficiently locate orphaned wells



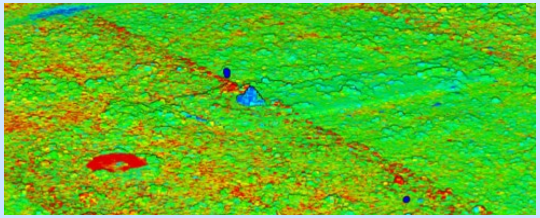
Magnetometer



Thermal

Overhead imagery

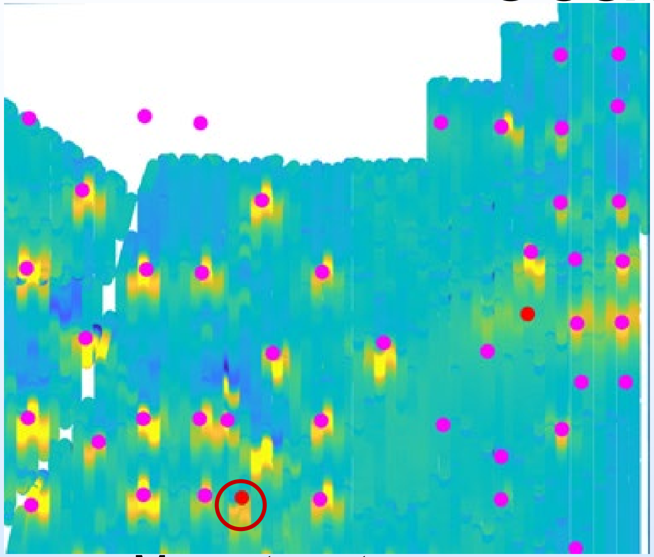
Lidar w/RGB coloring



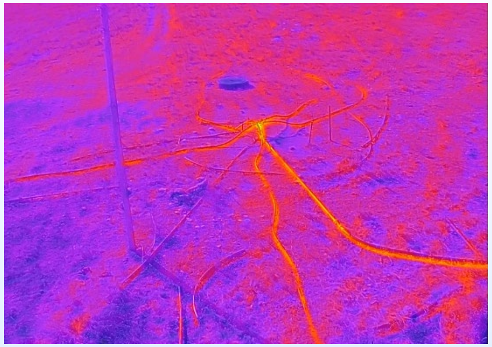
Lidar reflectivity

Topo map well with well head clearly seen in imagery but not in magnetometer

Good example of multi sensors able to locate orphaned wells



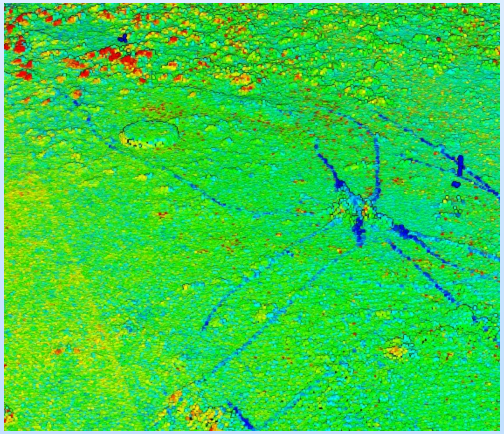
Magnetometer



Thermal

Overhead imagery

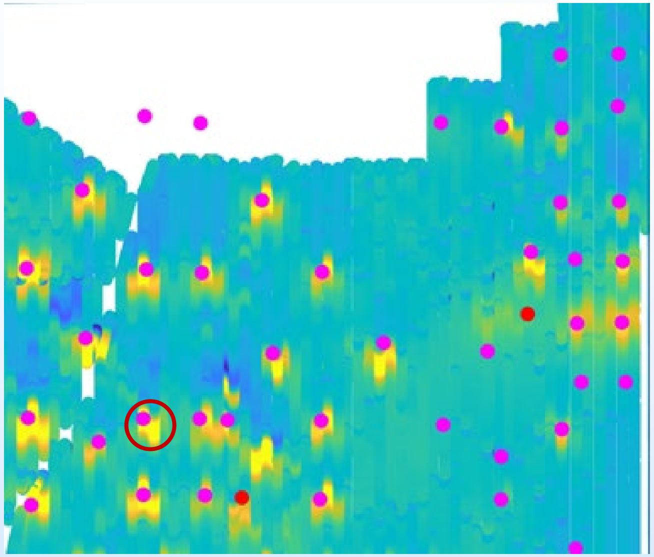
Lidar w/RGB coloring



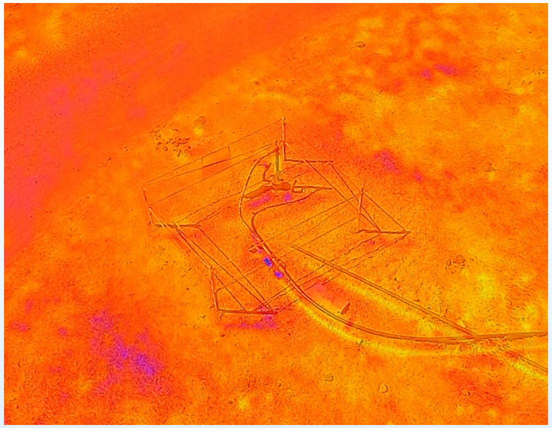
Lidar reflectivity

Topo map well with well head clearly seen in imagery and magnetometer, support infrastructure also seen in thermal and lidar reflectivity.

Example of active injection well



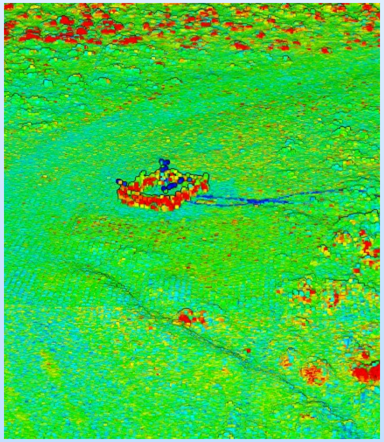
Magnetometer



Thermal

Overhead imagery

Lidar w/RGB coloring



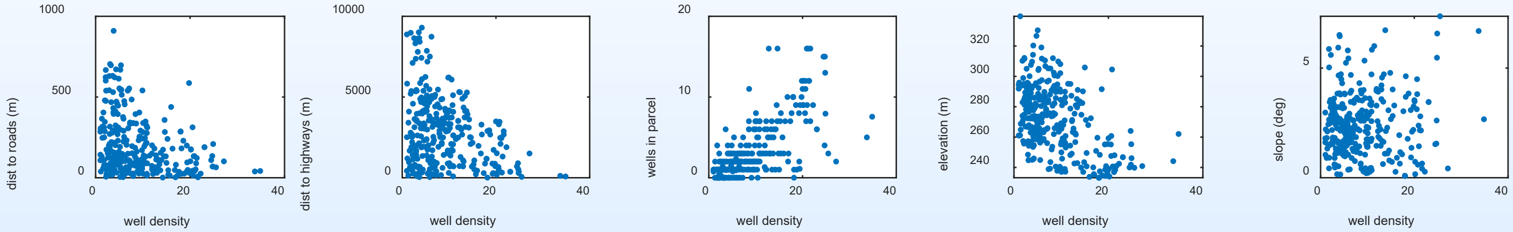
Lidar reflectivity

BIA database injection well with well head clearly seen in imagery and magnetometer
Well not seen in thermal imagery due to lack of thermal contrast

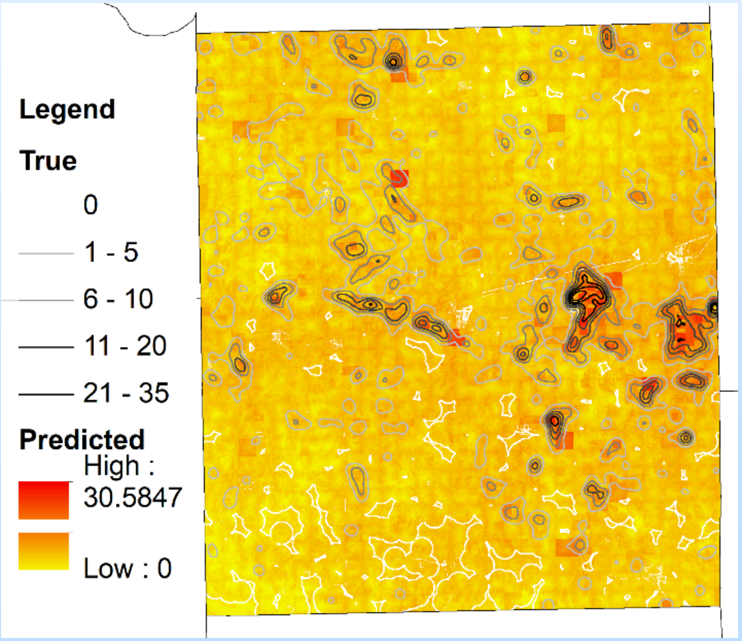
SNL – Integrated data products for analysis and ML to identify wells

Predictive Modeling- Linear Regression Approach

- SNL identified environmental factors indicative of O&G well density. Distance to roads and highways, slope, elevation, and land parcel well count all have moderate linear relationships with well density.
- Models* were only moderately successful in Lincoln County, OK.



* Results are similar for a ordinary least squares regression model and geographically weighted regression model.

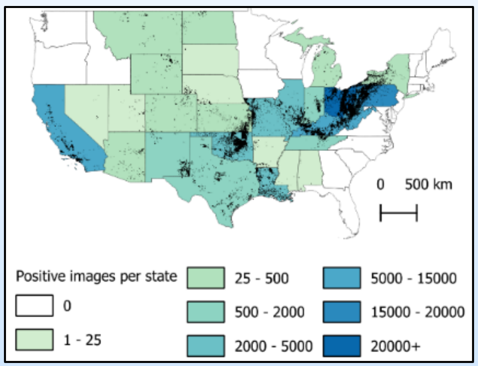


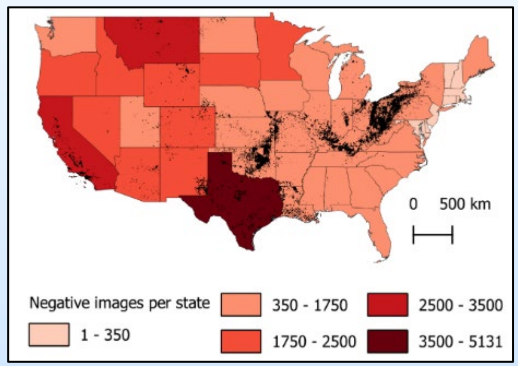
SNL – Integrated data products for analysis and ML to identify wells

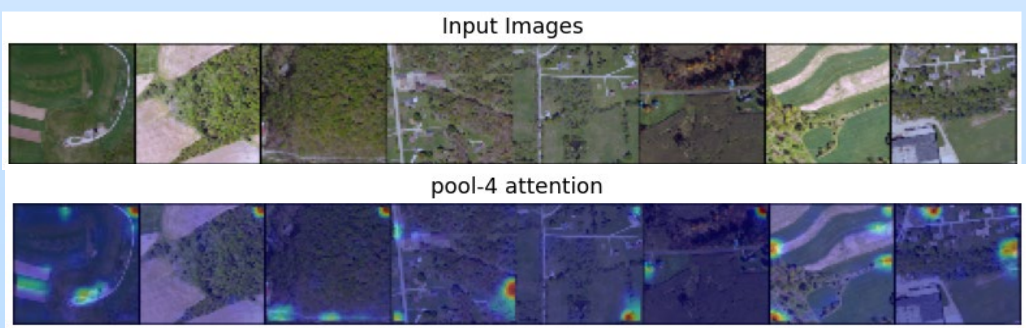
Classification with attention model

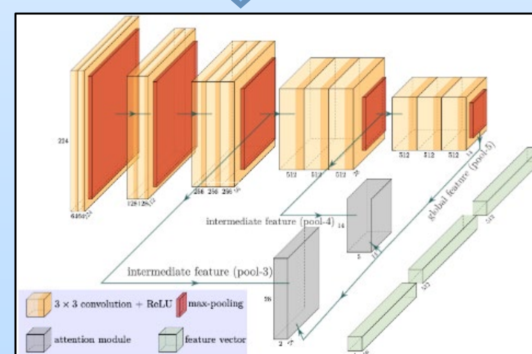
- SNL trained model with 2022 NAIP* imagery (n= 121k) and Boutot et al., (2022) documented orphaned well list.
- Attention model indicates areas in image responsible for classification
- Latest model now reaches 0.83 f1 score.

**National Agriculture Imagery Program*





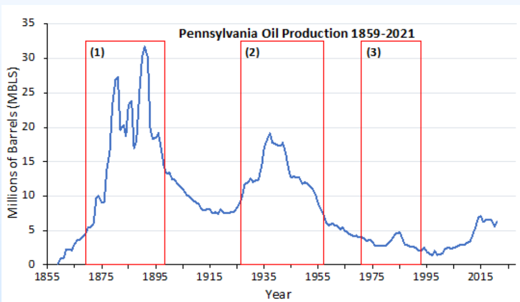


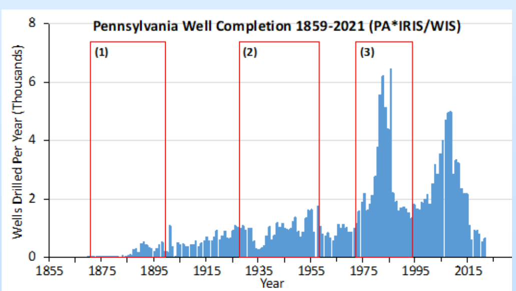


Adapted attention classifier from Yan et al. (2019) melanoma recognition via visual attention.

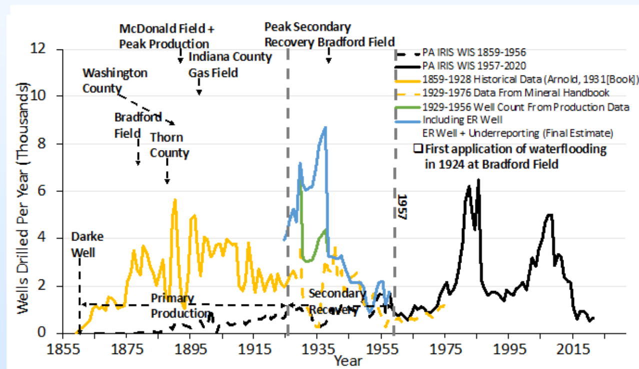
LANL- Identifying areas with high density of orphan wells using historical data and geographic information systems analysis

Input: Historical oil production data and number of wells drilled per year to find the irregularities in reporting

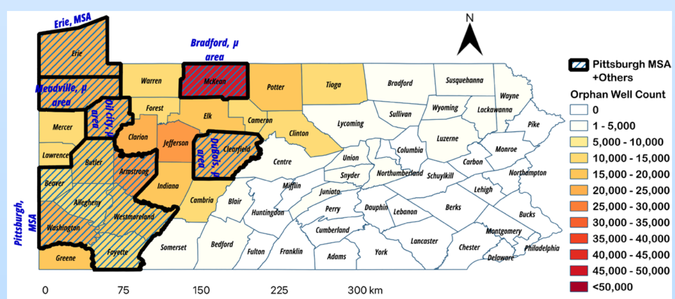




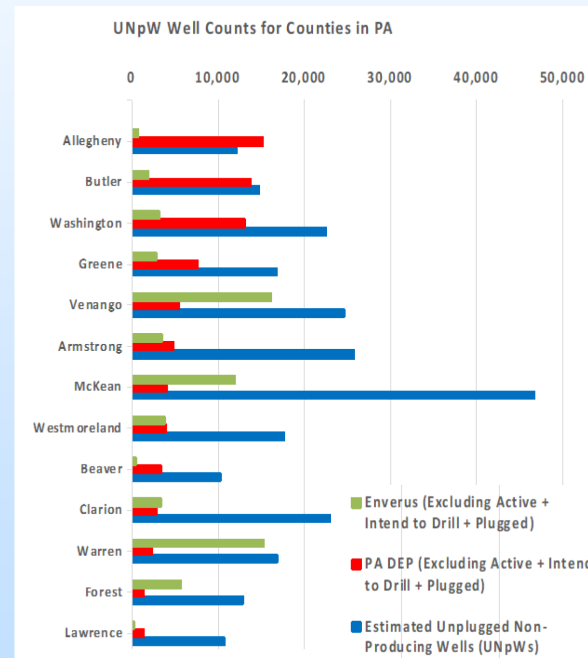
Estimation: Workflow based on boomtown study and statistical analysis to predict well counts. UOWs = Total well count in the database - Estimated well counts



Spatial Distribution: Estimated UOWs are distributed to counties based on the ratio found in the historical records



Field verification: Comparison of Unplugged non-producing wells (UNpWs) with PA DEP and Enverus databases. UNpWs = UOWs + DOWs+ Unplugged Abandoned wells

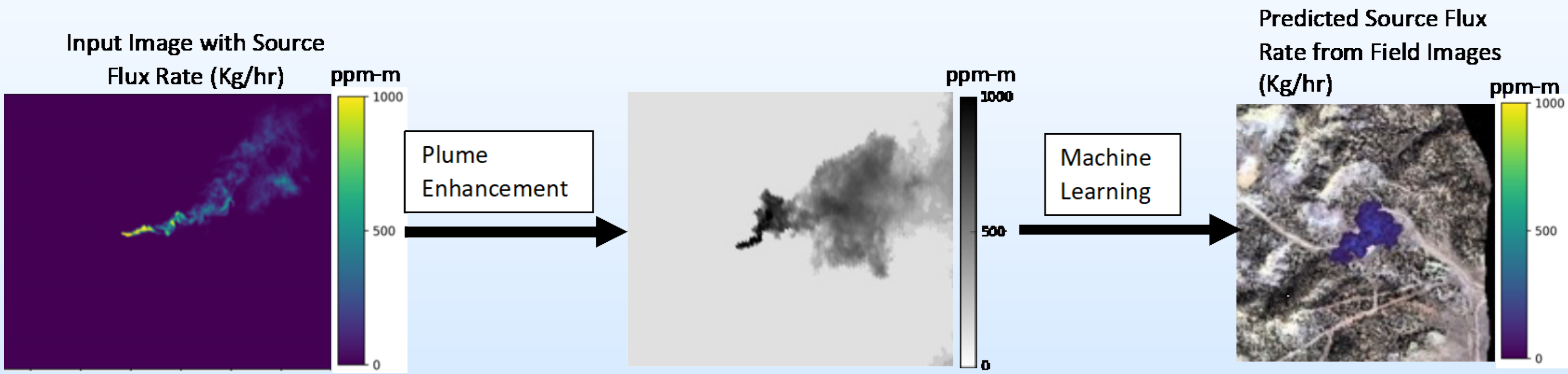


Using this approach, the estimated UOWs for PA is 340,827 and 309,462 in Oklahoma which is 3-fold higher than OK's estimated UOWs

LLNL-PRES-862255

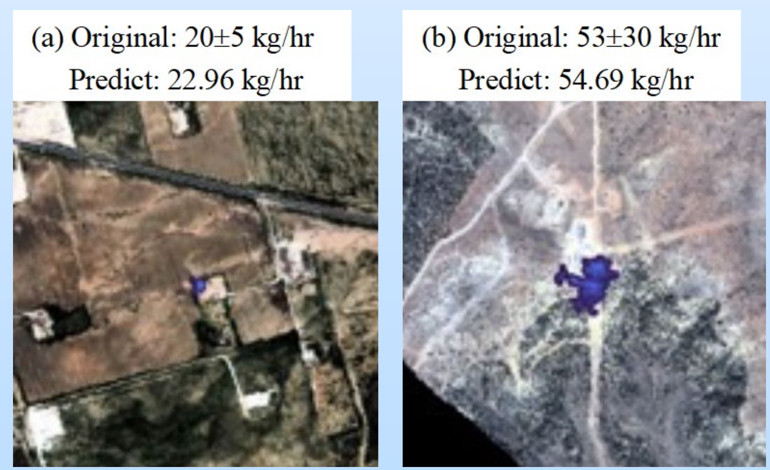
LANL - Quantifying methane emissions from infrared 2D images

Background: The airborne surveys can cover large areas, but the estimation of source flux rates from those surveys has 20-50% error using conventional source flux methods



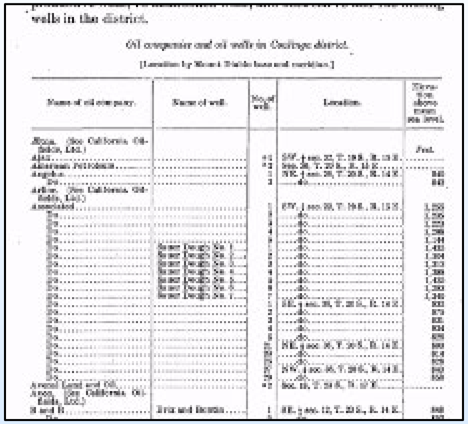
- We trained a deep learning model based on the synthetic airborne infrared spectrometer data which has 5% error in predicting source flux rates without any background information

- This ML tool can prioritize the leaky wells for sealing in large geographical areas without on-site surveys



Identify potential undocumented orphaned wells from Coalinga oil field report

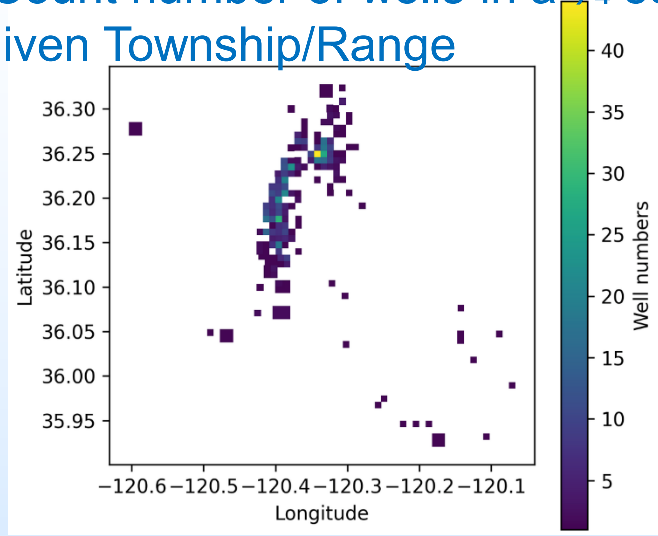
1. Mine well locations from tables in the report


wells in the district.

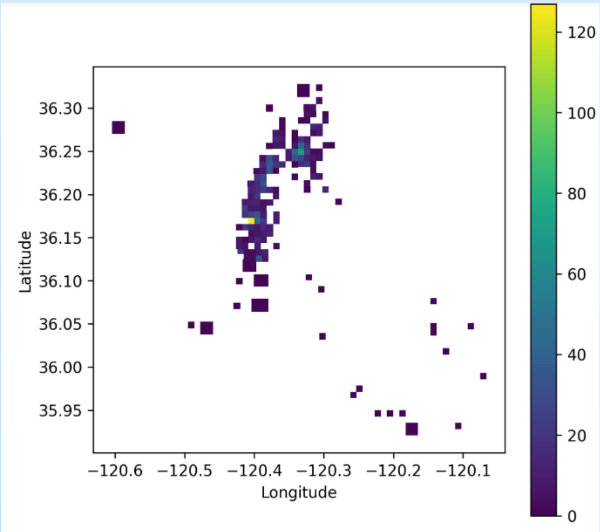
Oil companies and oil wells in Coalinga district.

Name of oil company.	Name of well.	No. of well.	Location.	Elevation above mean sea level.

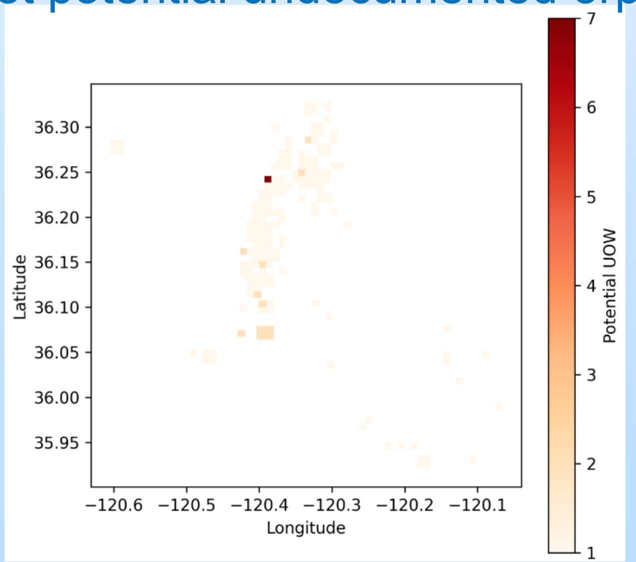
2. Count number of wells in a ¼ section of a given Township/Range



3. Compare with CalGEM database

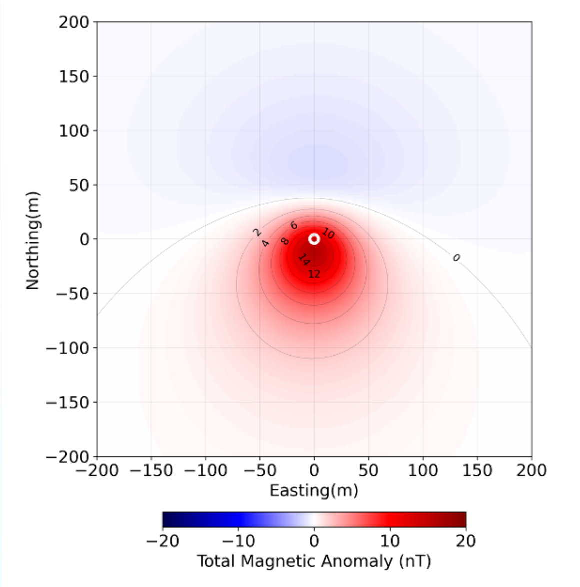


4. Get potential undocumented orphaned wells



Finding undocumented orphaned wells using magnetic anomaly data

- **Background**: The magnetic method has been successful in locating wells with steel casing. We developed a modeling program to estimate magnetic anomalies of a steel casing and generate synthetic data for training deep learning inversion model.



site: Osage Nation, OK

casing axis: Vertical

casing length: 200 m

OD: 0.254 m

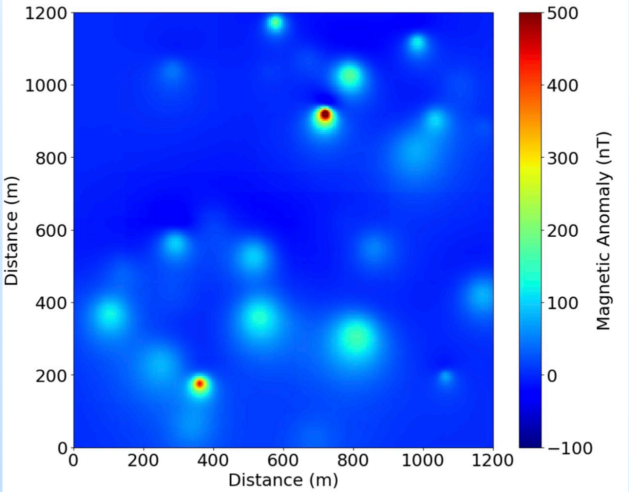
Wall thickness: 0.01 m

χ_m: 1000 SI

Sensor height AGL: 40 m

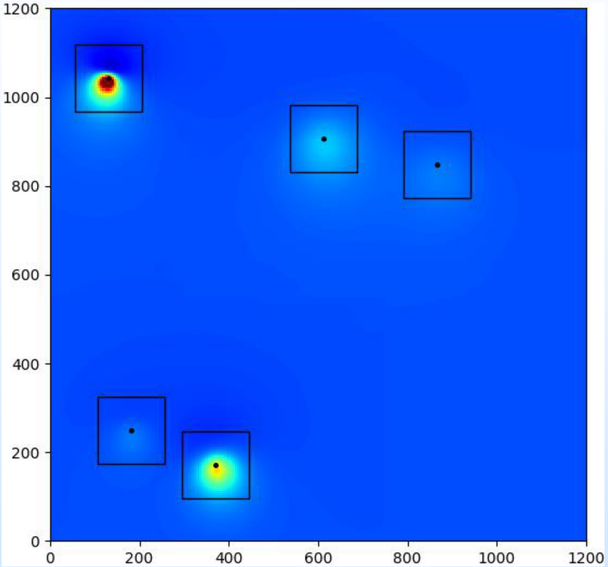
Magnetic anormal simulation of one orphaned well steel casing. Well parameters are on the right.

We generated 1000 forward simulation using random parameters.

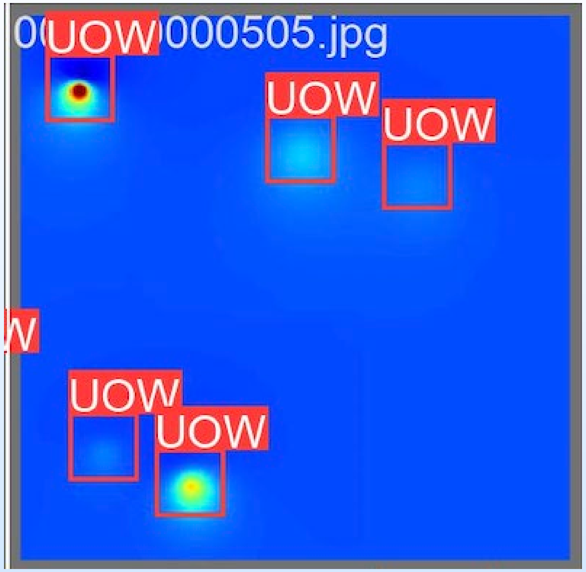


A random magnetic anomaly map of multiple orphaned wells generated from our magnetic modeling code.

Approach and current progress



Labeled magnetic anomaly for the validation of the model. Black box are desired magnetic anomaly region. Black dots are the well location

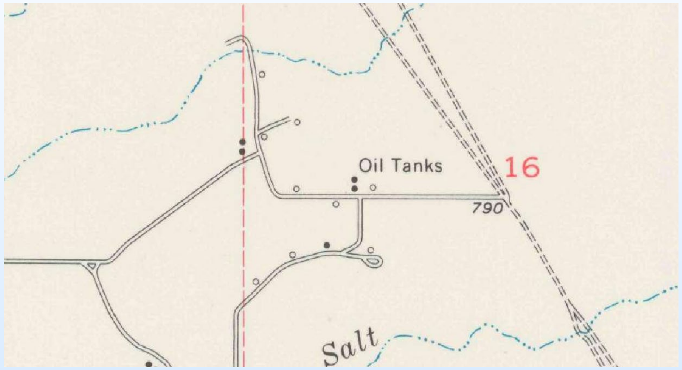


The orphaned well location detected by YOLO model. The red boxes are generated directly from the model.

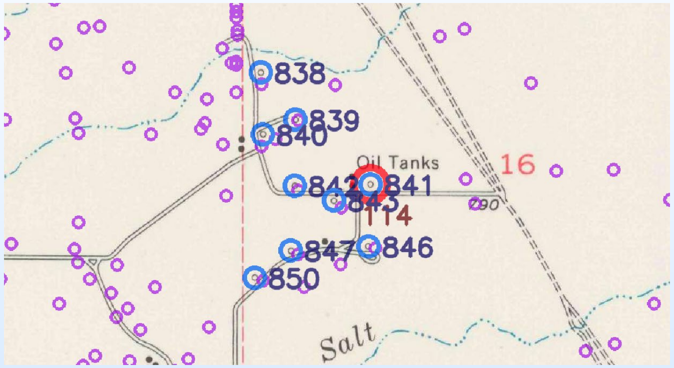
- We trained a deep learning model based on the popular object detection model YOLO (Redmon et al. 2016)to detect magnetic anomaly of wells.

Discovery of undocumented orphaned wells from historical topographic maps

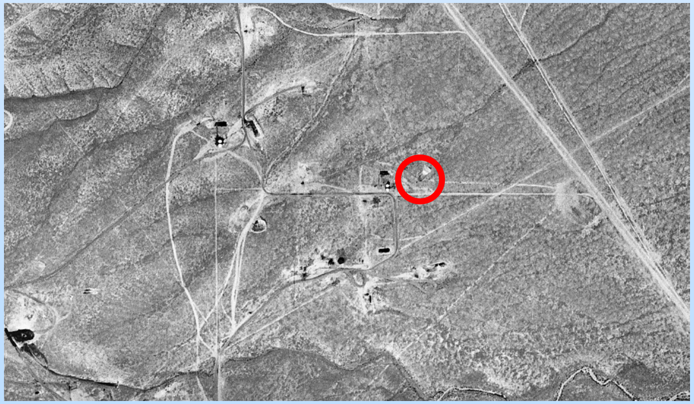
Input: inset of historical topographic map of the Belrigdge aera in Kern County, CA



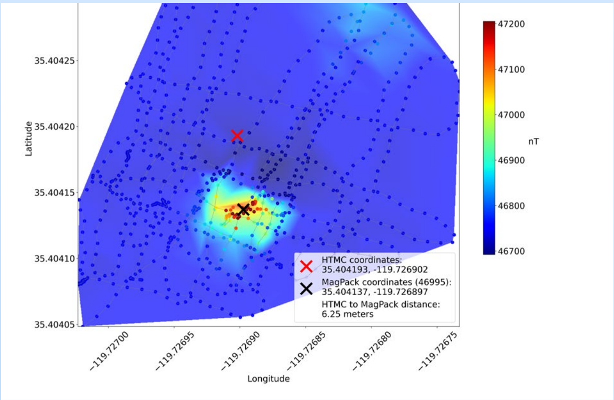
Detection: algorithm identified wells (blue circles), state documented wells (purple) and potential UOW (red)



Remote verification: Historical aerial map (1956) showing oil extraction activity and UOW



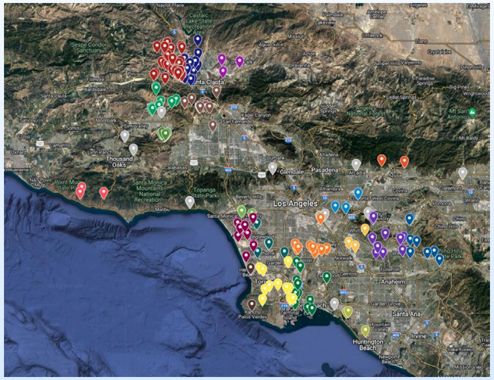
Field verification: Magnetic survey of UOW site displaying the presence of the UOW



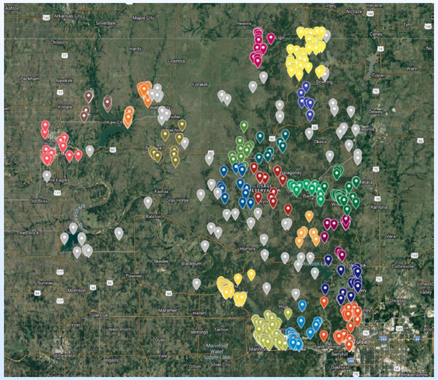
Discovery of undocumented orphaned wells from historical topographic maps

Found **1,301 potential UOW in 4 counties**: Kern, CA; Los Angeles, CA; Osage, OK; Oklahoma, OK

237 potential UOWs found in the Los Angeles County



487potential UOWs found in the Osage County area



From **magnetic survey** of **15 potential UOW**, the **average distance** between the topomap detected locations and field verified ones is **11.7±1.8 m**

Some potential UOWs have clear current surface expression

Summary Slide

a. Historical documents, particularly topographic maps are a good source of information for locations of potential UOWs, although field verification is essential to confirm their presence
b. Various estimates of UOWs through different methods. These need to be compared and verified
c. Collected a multi-sensor set of data at Osage Nation.
    a. Investigating use cases for each of the different sensors
    b. Future work will be processing data and looking at multi-sensor machine learning techniques
    c. Continue collecting data sets at different locations to build training set

Appendix

– These slides will not be discussed during the presentation but are mandatory.

Organization Chart

Work Package 3 Participants

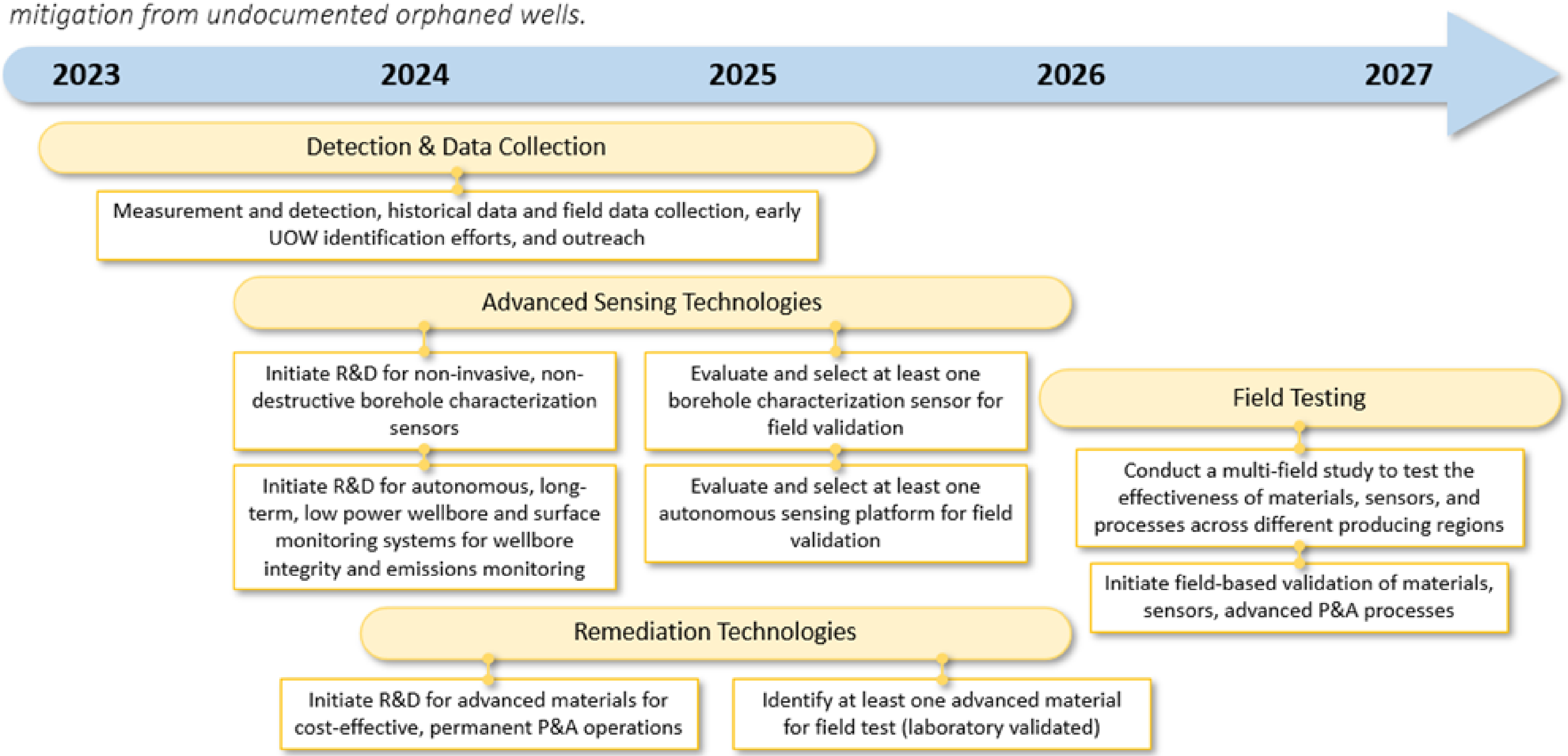
– Lead: Jacob Trueblood

– Co-Lead: Charu Varadharajan

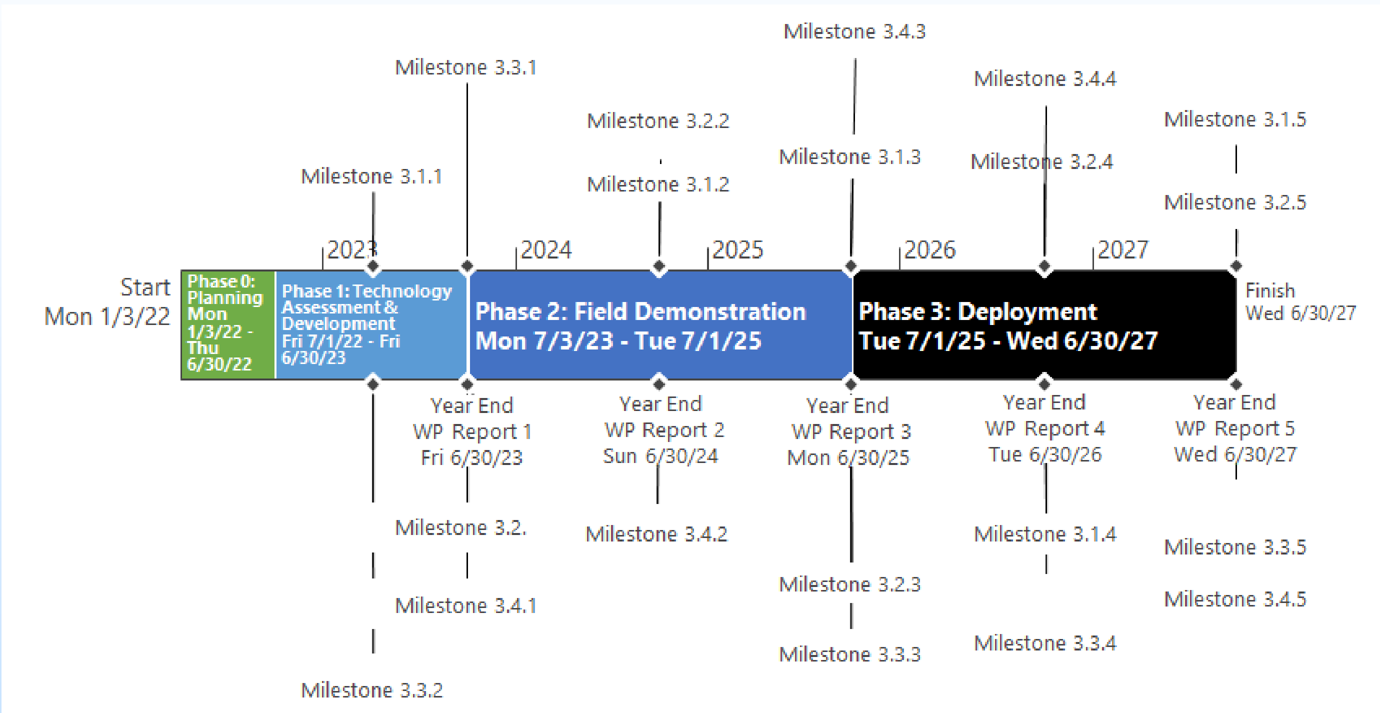
LLNL	LANL	SNL	LBNL	NETL
Jacob Trueblood	Dan O’Malley	Christine Downs	Fabio Ciulla	Jennifer Bauer
Ken Enstrom	Hari Viswanathan	Holly Eagleston	Charu Varadharajan	Mumbi Mundia-Howe
Xianjin Yang	Mohamed Mehana	Teeratorn “Meen” Kadeethum		
Yuan Tian	Javier Santos			
Hewei Tang	Anastasiia Kim			
Jaisree Iyer	Ismot Jahan			
Claire Knight				
James Reimer				

Gantt Chart (CATALOG program)

Overall Objective: *Develop advanced characterization, P&A materials, and long-term monitoring solutions for permanent emissions mitigation from undocumented orphaned wells.*



Gantt Chart (WP3)



- **Milestone 3.1.1** Initial technology assessment and requirements for Sensor Fusion, Data Integration, ML
- **Milestone 3.2.1** Integrate initial multi-sensor payloads, processing systems, and vehicular platforms
- **Milestone 3.3.1** Define data products, formats, storage/ingest, initial labeling
- **Milestone 3.4.1** Define initial framework architecture and off-the-shelf algorithm use
- **Deliverable 3.1** Lead/synthesize report leading to or in the form of a journal or conference paper with appendices fulfilling DOE needs and slide deck based on results of Year 1 activities. Support system demonstration and data collection.