Journal of Engineering Research and Reports

Volume 25, Issue 7, Page 18-31, 2023; Article no.JERR.102508
ISSN: 2582-2926

Application of a Machine-Aided Technique for Instantaneous Gas Leak Detection: A Case Study of Real-Time Modeling for JK-52 Gas Processing Plant

Godsday Idanegbe Usiabulu [a*], Ogbonna Joel [a,b], Livinus Nosike [a,c], Emeka Okafor [a,b] and Victor Aimikhe [a,b]

[a] African Center of Excellence, Center for Oilfield Chemicals and Research, University of Port Harcourt, Port Harcourt, Nigeria.
[b] Department of Petroleum and Gas Engineering, University of Port Harcourt, Port Harcourt, Nigeria.
[c] Integrated Elvee Services Ltd., Enugu, Nigeria.

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JERR/2023/v25i7934

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/102508

Original Research Article

Received: 02/05/2023
Accepted: 10/07/2023
Published: 27/07/2023

ABSTRACT

Many fluid leak detection mechanisms rely on observation of volume changes and physical evidence of leak, which may take hours, days and sometimes weeks or months to be seen. This is a concern in gas plants where the proximity of the leakage may constitute environmental pollution as well as health hazards for personnel in the vicinity. Economic losses have also resulted from delays in mitigating a gas leak problem due to late detection.
This study applies a machine learning technique to develop an algorithm that can detect gas leak in

**Corresponding author: Email: godsdayusiabulu@gmail.com;*

real-time, where the only possible delay is the lag-time between the inlet gauges at the upstream valve and the outlet gauge at the downstream valve. In this case study of JK-52 gas processing plant, the difference pressure gauge readings were calibrated against the volume of the gas in the inlet section to quantify the leak volume. Because gaseous fluids do not present physical indication of volume, a pressure-based method was used for the detection, where drop in gauge pressure due to de-pressurisation indicate leakage in the absence of recorded gas supply or collection.
Python coding language, using Jupyter and Pycharm Integrated Development Environments (IDEs), was used for the programming. The machine learning algorithm analyses the incoming streaming pressure versus time datasets from the gauges during the residual and ramp-up flow phases to set the acceptable pressure difference cut-off. A minimun difference in gauge reading may be normal within an acceptable error margin. The change in the consistency of reading within this acceptable window defines the tolerance. The system is set-up to blare an alarm when there is leakage, usually based on a cut-off or tolerance, to be detected by the machine-aided process.

Keywords: Machine aided; application; gas leak; python; coding; Jupyter; pycharm integrated development environments.

1. INTRODUCTION

"Natural gas is composed mostly of methane, the simplest hydrocarbon molecule, with only one carbon atom. But most gas at the wellhead contains other hydrocarbon molecules known as natural gas liquids, such as ethane (with two carbon atoms) and propane (with three carbon atoms). Therefore, it is sent to processing facilities, where most of the natural gas liquids are removed and sold separately" [1-3]. "Gas processing facilities produces consumer-grade natural gas, which is primarily made up of about 95% methane. During these operational services, gas leaks tend to occur undetected early enough. This undetected gas leakage can lead to undesirable economic loss of natural gas from installed facilities and are often accompanied by toxic air pollutants that typically pose safety and public health concern" [4-7]. However, for safety officers and plant managers trying to keep up with the evolution of detection technology are finding it difficult since no single system or technology is the solution to every plant's problem (Tan & Tan, 2019). This research will develop a model using machine learning algorithms to detect gas leaks based on available data and present the one with highest accuracy.

1.1 Research Problem

Gas leaks are invisible, unregulated and the majority of them go unnoticed. These leaks may depend on any of the following: operation practices, equipment age and maintenance. Leaks and venting at natural gas processing plants release other pollutants (e.g., benzene, hexane, hydrogen sulfide) besides methane that can threaten air quality and public health [8-11]. Hence, there is need for early detection of gas leaks by using appropriate Machine Learning models. Nigeria is a province of gas with pockets of oil, and the use of pipeline is considered as a major means of conveying petroleum products which serves as the major assets to the Nigeria economy and should be well protected.

1.2 Research Objectives

This research work is aimed at modeling innovative technological leak detection for gas processing plants using machine learning.

The objectives are:

1. To develop a model for gas leak detection using coding and machine learning.
2. To validate the model sample gas plant data.
3. To generate a safety notification and improved flow system based on leak detection findings.
4. To create a programme or coding package for identifying leaks and notification, as a means of curtailment in gas leak prone flow systems.

2. DATA AND METHODOLOGY

The methodology that was followed during this study includes important steps to building a Machine learning model. The first step was to collect the required dataset and a preprocessing phase which includes cleaning the data, attempting a linear regression model and other regressions (Random Forest). The linear regression model and the random forest were used for predicting tolerance and gas leak detection.

The second step is to train the proposed model and evaluate its performance. A detailed description of the methodology is included in this paper. Fig. 1 summarizes the methodology of this study.

2.1 Fluid Leaks Detection Mechanism

Natural gas is considered a fossil energy source from under the earth's surface. Its main component is methane, but it can contain non hydrocarbon gases and natural gas liquids as well. It comes from plants and animals that were initially in thick layers on ocean floors, the earth's surface, and other areas [12].

Leak detection is a method in which the existence of a leak within a system is determined. The techniques used here is utilized across a wide range of systems where a container must seal in some material.

Gas leak detection is thus the process of identifying potentially hazardous gas leaks by sensors.

Thermal imaging which is the most common gas plant leak detection method, was used. It detected natural gas leaks from pipelines due to the differences in temperature between the natural gas and the immediate surroundings. This method can be used from moving vehicles, helicopters or portable systems and is able to cover several miles or hundreds of miles of pipeline per day.

2.2 Physical Evidence through Observation

The oil and gas industry is stepping up efforts to find and repair methane leaks from natural gas pipelines.

The result showed that "Addressing methane leakage from pipelines has come into sharp focus". "This shift has been driven by learnings from new measurement campaigns and a growing need to reduce methane loss from the entire supply chain to mitigate climate change, improve carbon accounting and enable the demonstration of responsibly sourced gas."

When it is combusted, natural gas has a much lower carbon intensity than coal. "However, uncombusted methane has a global warming potential that far exceeds carbon dioxide–the Environmental Protection Agency estimates as much as 28-36 times more–and leaky supply chains can cancel or reverse the climate benefits of natural gas".

Identifying pipeline natural gas leaks, which are invisible and often odorless, is a significant challenge. "Legacy methods include walking along pipelines with hand-held instruments (e.g., organic vapor analyzers and combustible gas indicators) and flying aircraft along rights-of-way to search for visual signs of disturbance (e.g., dead vegetation and encroachment). Although legacy methods find leaks, their overall effectiveness remains unclear, and a growing

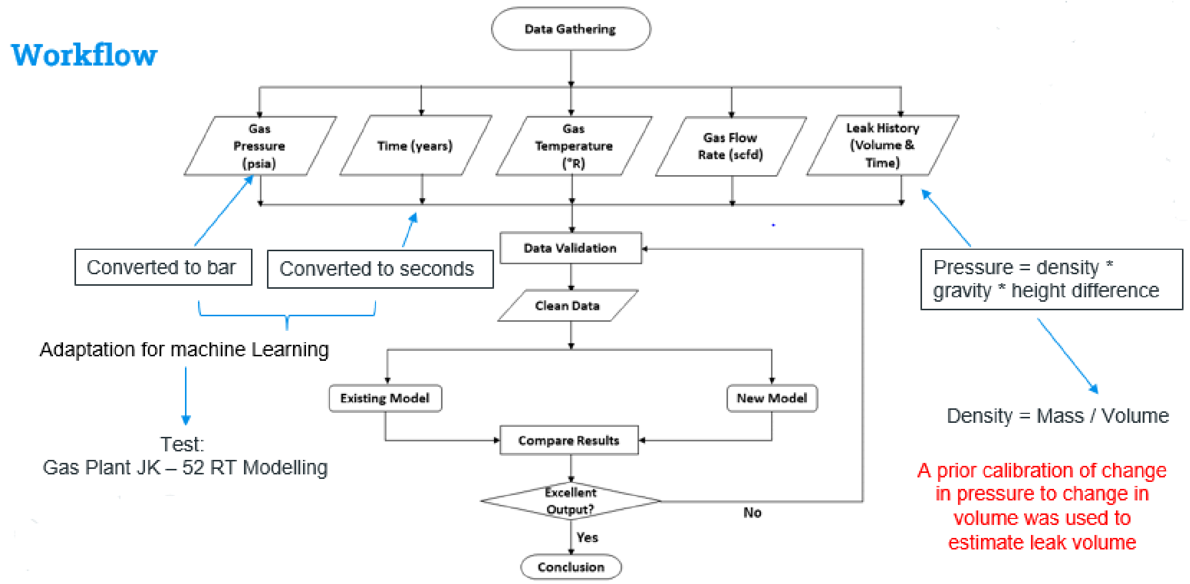


Fig. 1. Summary of the methodology

body of research demonstrates that pipeline methane emissions are of greater significance than previously thought."

Innovators are improving methane detection techniques and economics. In fact, the research showed that at least 100 distinct leak detection technologies are commercially available. Leak detection programs often begin by surveying wide area. Although technologies with broad coverage, such as satellites, planes and helicopters, generally entail higher detection thresholds, this report commends this approach.

"Rapidly screening sites for large sources to direct more targeted close-range surveys presents an important value proposition due to the emissions profiles typical of most oil and gas systems. Methane leak emission rates generally follow a highly skewed distribution in which a small number of leaks account for the majority of overall emissions. Therefore, it can be concluded that screening more frequently for large leaks–even at lower sensitivities–can be more effective than less frequent close-range inspections.

2.3 Volume Changes

Mass (or volume) balance leak detection methods follow the principle that the metered inlet volume, less the metered outlet volume, less the change in mass inventory (or line pack) due to the compressibility of the fluid and pipes should always be zero if the pipeline is not leaking, this is

$$V_L = V_{in} - V_{out} - \Delta V$$

where

V_L = leakage volume;
V_{in} = metered inlet flow;
V_{out} = metered outlet flow and
ΔV = pipeline pack or inventory.

This method and the pressure-flow deviation method are the most used model-based leak detection programs in the oil and gas industries (Griebenow and Mears 1989; Furness and Reet 1998).

I suggest the combination of both methods to give a very good Leak Detection, since the use of one method doesn't interfere with the other.

2.4 Aerial Photographs

One of the traditional screening methods involves flying over a pipeline right-of-way in search of visual pipeline failure indicators, such as dead vegetation, ground disturbances, melted snow and right-of-way encroachment. Usually, areas of concern would be flagged for follow-up using handheld methods.

Today, aircraft can be equipped with devices that detect methane, such as infrared spectrometers and optical gas imaging cameras. "Aircraft detection limits range from a few kilograms of methane per hour to dozens of kilograms per hour. "This technology is readily available for deployment and has undergone multiple third-party controlled release tests to verify performance metrics for aboveground infrastructure.

"The main advantage of aircraft technologies is the more significant spatial scale, providing the ability to survey hundreds of miles of pipeline per day or, depending on the infrastructure density, hundreds of sites per day.

2.5 Gas Plants Leakage Proximity Concerns

"A natural gas leak refers to an unintended leak of natural gas or another gaseous product from a pipeline or other containment into any area where the gas should not be present" (Awad, et al*.,* 2020). "Gas leaks can be hazardous to health as well as the environment. Gas leaks from pipelines may give an odour of gas in the air while gas from landfills may not give an indication of odour. Affected soil from a gas leak will have a characteristic blue-black colour and rotten egg odour. Roots killed by gas will be blackened and necrotic" (Baker, 2002).

"Natural gas leaks can also cause smaller-than-normal leaves on trees, wilted plants and yellowish patches of grass. Symptoms of exposure (Physical symptoms of natural gas poisoning) to low levels of natural gas include headaches, dizziness, fatigue, nausea and irregular breathing" [4]. "The most common cause of gas leaks is damage to underground utility lines. If you will be digging on your property, does it safely to avoid breaking gas utility lines (as well as other utilities like fiber-optic cables)" [4].

"Even a small leak into a building or other confined space may gradually build up an explosive or lethal concentration of gas. Leaks of natural gas and refrigerant gas into the atmosphere are especially harmful due to their global warming potential and ozone depletion potential" [4]. Leaks of gases associated with industrial operations and equipment are also generally known as fugitive emissions.

2.6 Environmental Pollution

"Leaks are considered very dangerous since they can build into an explosive concentration. They can kill vegetation and trees, cause explosions and fires, and might release greenhouse gases into the atmosphere. When you're looking for gas leak signs in your home, keep in mind that it might not have physical signs or smells. You can find a damaged gas pipe, dead houseplants, and if a smell is present, it'll be rotten eggs and sulfur. Near the gas line, you might see a white or dust cloud, and a whistling or hissing sound" [12].

"In addition to wasting a source of energy, leaked natural gas mostly methane is a powerful greenhouse gas. It is a significant contributor to climate change that makes it essential for gas utilities, and the regulators and public officials that oversee them, to act swiftly and decisively to repair and prevent all methane leaks. The gas utilities' pipe systems are just one link in the national gas supply chain that brings gas from the well to your home. Leaks are an issue at every stage, starting at the wellhead. That's why we're addressing leaks throughout the system" [13].

2.7 Health Hazards and Personnel Safety

Most leaks don't pose an immediate threat to safety, but some can [14]. "We have shared the maps and leak indicators with local gas companies. If you ever smell gas, or have any reason to suspect a problem, immediately exit the building or area. Don't light matches or smoke, and don't use any electrical devices, including a phone, until you are away from the suspected leak. Then, call your local utility. The major health concern about outdoor methane leaks is that they contribute to smog, which aggravates asthma and other respiratory conditions" [14].

2.8 Impact of Detection Delay on Mitigation Strategies

The impact of detection delay on mitigation strategies are:

a. It leads delay in instrument response and could lead to a major accident.
b. It could also lead to failure in protective device to operate on demand.
c. There will be increase in unavailability of the system thereby rendering the system ineffective.
d. This will also lead to reduction of output.

2.9 A New Machine Learning Technique for Gas Leaks Detection

The technique used in this research is in line with a new study of a natural-gas leak-detection tool pioneered by Los Alamos National Laboratory scientists that uses sensors and machine learning to locate leak points at oil and gas fields.

Automated leak location system finds gas leaks fast, including small ones from failing infrastructure, and lowers cost as current methods to fix gas leaks are labor intensive, expensive and slow. Our sensors outperformed competing techniques in sensitivity to detecting methane and ethane. In addition, our neural network can be coupled to any sensor, which makes our tool very powerful and will enable market penetration.

2.10 Algorithm for Real-Time Detection

"YOLO (You only look once) is a new algorithm which means that an image can predict the objects and their locations at one glance. It uses neural networks for real-time object detection. This algorithm has evolved over the years, it started with YOLO v1 (or unified) – It has several localization errors, Yolo v2, YOLO v3, YOLO v4. Currently, YOLO v3 is the state of art algorithm which is used for single stage object detection. YOLO v3 can basically achieve its real-time performance on a standard computer with graphics processing unit (GPU)" [4]. "The whole framework only needs to use a relatively simple structure of Convolutional Neural Network (CNN) to directly complete the regression of target detection to predict the position of the bounding box and the class of the candidate box" [12]. "YOLO focuses on the entire image as a whole and predicts the bounding boxes and then calculates the class probability to label the boxes. It predicts limited number of bounding boxes to achieve its goals. It can classify objects up to 155 FPS (frames per second) in real time, achieving twice the mean average precision (mAP) of other object classifiers. It is a single convolutional network that simultaneously predicts multiple bounding boxes on multiple objects and then generates a class probability for that object" [4].

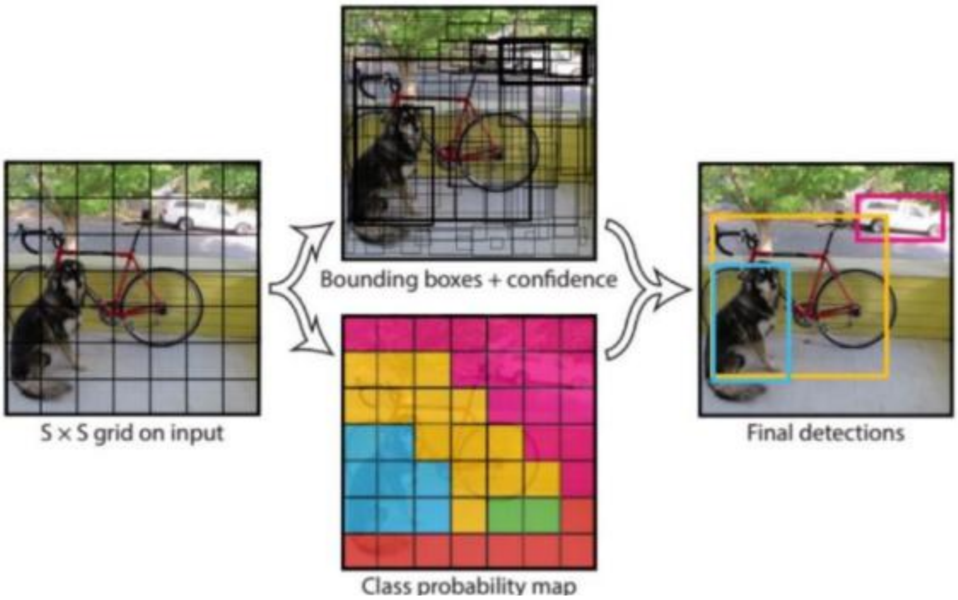


Fig. 2. YOLO bounding box, object detection and localization

In YOLO: -

- The image is divided into M grids, each grid having equal dimensional regions P × P. Each of these grids are responsible for detecting and locating the objects present in it.
- These M grids predict their bounding box coordinates relative to the cell coordinates, along with the object label and the probability of it being present in the cell.
- This highly decreases the computation rate as the cells of the image handle both detection and recognition.
- Non-Max suppression is used to filter through all the boxes, and also eliminates overlapping boxes and duplicate predictions.

The above method can only detect object. This research is based on detecting gases and so a new approach is proposed for gas volume estimation in the JK-52 real-time gas modeling case. The added value of this current gas leak detection modeling is that is based on a *normalization* of pressure flow, unlike volume changes that presents with many limitations.

2.10.1 Normalization algorithm

$$X_{normalised} = \frac{(X - \text{MinValue})}{(\text{MaxValue} - \text{MinValue})}$$

where

X represent a data point

Normalisation present every data in terms of percentage or fraction with respect to its distribution

In this case, equivalent pressure and volume data are each presented in percentage or fraction, so they are comparable.

2.11 Short Lag-Time Delay Between Gauges

Drilling personnel need to fully understand about the lag time. Lag time is time delay from pressure adjustment made on a choke valve or a choke Hierarchical classification and regression HCR to show up on the drill pipe pressure gauge.

When you adjust a choke position, you will not be able to see changes on the drill pipe gauge right away because drilling fluid is compressible so you need to be patient and wait a little bit until you see the changes on the drill pipe gauge.

2.11.1 How can we know how many seconds for the lag time?

We can roughly estimate the lag time about 2 seconds. This number is just a rule of thumb. If you want the actual lag time, you can determine it by performing a choke drill. You really need to know about the lag time otherwise you can get confused a lot when you attempt to adjust drill pipe pressure.

2.12 Application to JK-52 Gas Processing Plant

In the gas plant of study, the effluent (a mix up of water, oil and gas) is pumped into the gas plant from nearby oil well.

Crude stored in a Floating Production Storage and Offloading Offshore may also tapped from a Tanker offloading / lifting buoy and transported to the Gas Plant.

There is also provision for piped crude from multiple well clusters in the field to ensure constant source of hydrocarbon.

The crude passes through a water-oil-and-gas separator, a purifier or a compressor as part of the refining or treatment process before delivering the final gas product.

2.13 Pressure-Based Method for Gas Detection

- The first step of the modelling is to plot the pressure readings of the inlet (orange colour) and outlet gauge (blue colour) in a Pressure versus Time chart.
- The second step is to carry out analytics. Note is taken to not confuse recorded pressure changes due to gas sampling / lifting / sales, etc, as leakage at the observatory or display or data displace screen which may be at base in town
- Machine Learning helps to automate the plots, so that the analytics (Tolerance, Lag and Consistency) are carried out in real-time (once the initial lag-time has elapsed), as shown in the Animation.

2.14 Python Programming Languages Coding

The following are the steps taken:

- python IDEs (Jupyter and Pycharm were used)
- Among the libraries used are: Pandas, Numpy and Matplotlib
- *PiP* was used to install *SciencePlots*
- The Input Data set was set as a .CSV file for the initial coding
- The sample data was used to *train* machine on tolerance
- The data was *cleaned-off* for artefacts before plotting on Python development environment
- The *visualisation* was used to define the parameters of display including colour and labelling
- The early flow stages when there were still *residual gases* in the system and when flow was *ramped up* were used to determine the tolerance
- This tolerance will vary and *machine learning* helps to determine it with different and more incoming data
- The available data is *split* to train the machine and build *regression models* which ware tested on the *training dataset*
- The best regression, in this case *random forest*, provided the most accurate result and is retained for the given case study

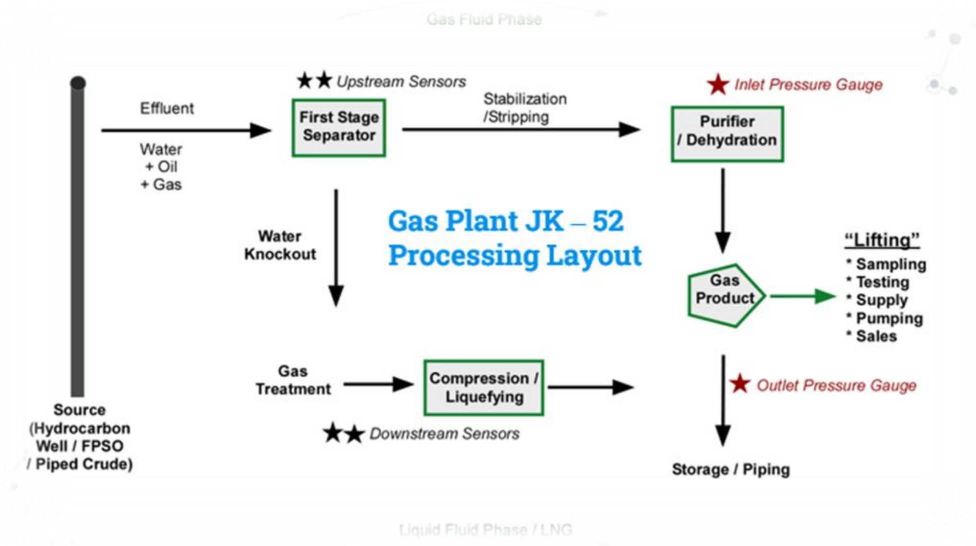


Fig. 3. Summary of process

FPSO: Floating Production Storage and Offloading

- The test score between the *training accuracy* and the *test accuracy* is shown to confirm prediction
- The process is *automated* to work in *real-time* and the *animation* generated for the presentation purposes
- This involves importing and running the useful *libraries* and *plotting styles* to show the arrays and follow the sequences of the analytics
- These were used for *repartition* and *enumeration* of the animation, which appended the plot parameters including colour and labelling
- The annotation for gas collection or lifting is set as is different for that of leakage, where leak is indicated when pressure drop is *eventless* / or *causeless*
- Expected streaming or real-time data is set to trigger *colour code alarm* in the system when leak occurs.

The codes are presented in the appendix.

2.15 Use of Different Ides and Regressions

In a world where nearly all manual tasks are being automated, the definition of manual is changing. There are now many different types of Machine Learning algorithms, some of which can help computers play chess, perform surgeries, and get smarter and more personal.

We are living in an era of constant technological progress, and looking at how computing has advanced over the years, we can predict what's to come in the days ahead.

One of the main features of this revolution that stands out is how computing tools and techniques have been democratized. Data scientists have built sophisticated data-crunching machines in the last 5 years by seamlessly executing advanced techniques. The results have been astounding.

The many different types of machine learning algorithms have been designed in such dynamic times to help solve real-world complex problems. The machine learning algorithms are automated and self-modifying to continue improving over time. Before we delve into the top 10 machine learning algorithms you should know, let's take a look at the different types of machine learning algorithms and how they are classified.

Machine learning algorithms are classified into 4 types:

- Supervised
- Unsupervised Learning
- Semi-supervised Learning
- Reinforcement Learning

2.16 Analyses of Pressure Versus Time Datasets

Here are the steps carried out to analyze time series:

1. Data were collected and cleaned.
2. Visualization with respect to time vs. key feature was prepared.
3. The stationarity of the series was observed.
4. Charts were developed to understand its nature.
5. The model – Autoregressive (AR), Moving Average (MA), Autoregressive Moving Average (ARMA), and Autoregressive Integrated Moving Average (ARIMA) were built.
6. Extraction of insights from prediction.

2.17 Machine Leaning of the Real-Time Algorithm

Real-Time Machine Learning: It is the process of training a machine learning model by running live data through it, to continuously improve the model. This is in contrast to "traditional" machine learning, in which a data scientist builds the model with a batch of historical testing data in an offline mode.

Real-time machine learning is useful in scenarios when there is not enough data available upfront for training, and in cases where data needs to adapt to new patterns. For example, consumer tastes and preferences change over time, and an evolving, machine-learning-based product recommendation engine can adjust to those changes without a separate retraining effort. Therefore, real-time machine learning can provide a more immediate level of accuracy for companies and their customers by recognizing new patterns and adapting to reflect those.

2.18 Residual and Ramp-Up Flow Phases

- Process 1: The gas plant stabilises and strips lighter gas or condensates to produce purified dry gas ready as end product
- Process 2: The alternate process processes crude effluent by first separating the water and trace or associated oil,

before it is treated to remove impurities such as Carbon dioxide and sulphides). The resulting gas is then compressed or liquified (Liquified Natural Gas – LNG) for storage and eventual supply.

In both cases, initial sensors and gauges are placed at the upstream (sourcing section) and at the downstream (receiving section) of the products. Inlet and outlet pressure gauges are placed across intervals with tendency of gas leak.

2.19 Split and Training Datasets

Machine learning models were built after pre-processing and cleaning the dataset. To build this model, the dataset was divided into samples to train and test the model. The model's performance was measured in terms of accuracy on the testing sample. The most common approaches for splitting the dataset are 6:4 (training: testing). In the 6:4 approach, the dataset is divided into two samples, one for training and the other for testing. The training sample represents 60% of the dataset, and the testing sample is the remaining 40%. The training sample is used to train the model and enhance its ability to learn the complexity behind the features of the dataset, whereas to measure the performance of the model on unseen data the testing sample is used.

3. RESULTS AND DISCUSSION

The results obtained are shown below:

3.1 Determination of Alarm Cut-Off by Machine Aided Process

Alarm management is an effective solution to operation safety and efficiency in many industries. As the modern process industry becomes more complex and digitalized, alarm management becomes a necessity. However, the current status of alarm management applications is unsatisfactory, with too many alarms to convey valueless information to operators or even disturb them. Conventional alarm management can significantly alleviate alarm overloading but has difficulty in recognition and presentation of true abnormal situations. As an ultimate goal to generate one and only one alarm under a certain abnormal situation, we should meet and go beyond the standards and guidelines to achieve a smart alarm management. For this purpose, advanced alarm management should be developed and applied.

3.2 QC Using Consistency and Tolerance Windows

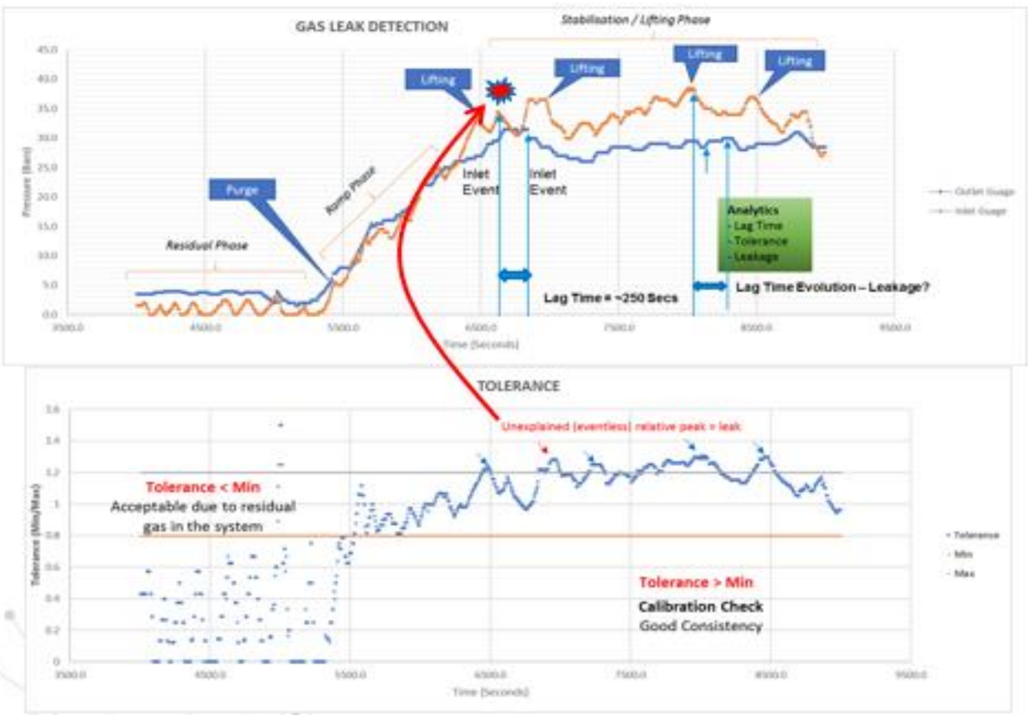


Fig. 4. Detection technics for machine learning

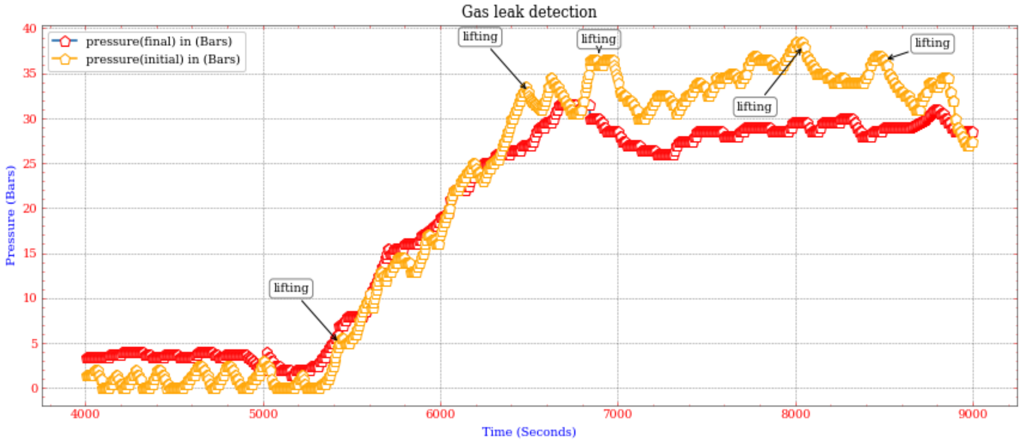


Fig. 5. Plotting with python

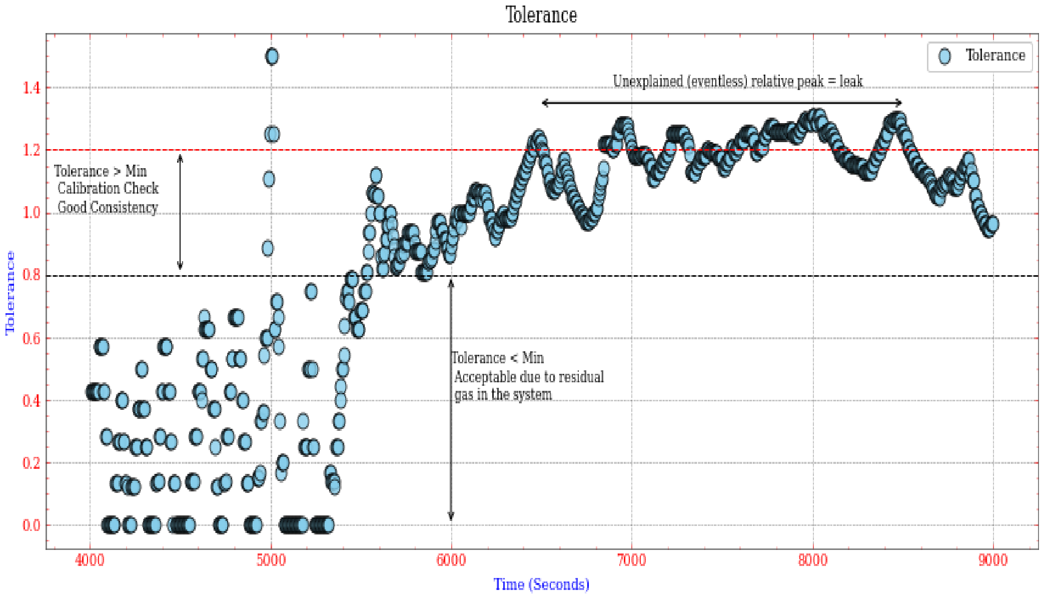


Fig. 6. Visualisation of tolerance

4. SUMMARY

4.1 The System Alarm Alerts on Leakage Detection

Gas leaks can cause significant damage and result in high costs for building owners, tenants, and property managers. That's why leak detection systems have become a crucial aspect of building management. In recent years, advancements in technology have made it possible to detect leaks automatically and remotely, thanks to machine learning algorithms.

Leak detection systems typically use machine learning algorithms to profile tenant behaviour using historical data. By analysing the volume and time of gas usage during a typical weekday or weekend, the algorithm can recognize events and predict future consumption. Using the data acquired, alarm thresholds are established based on past maximum consumption events. By splitting these events by the day of the week and further dividing them by time, the algorithm can accurately detect abnormal water usage patterns and trigger an alert if necessary.

5. CONCLUSION

Input gas data is calibrated and evaluated for consistency in real-time. The data is then corrected for lag and used to compute Tolerance.

Min. and Max. Tolerance Cut-Off is set based on machine training dataset. Where value is higher than maximum cut-off, machine sets off alarm. Time of alarm is checked against events such as lifting, residual gas, etc.

Where alarm is eventless, leak is suspected and eventually confirmed. Further modelling becomes predictive as machine learns from experience.

ACKNOWLEDGEMENT

Special thanks goes to Dr. Livinus Nosike, Head Consultant, Integrated Elvee Services Ltd, for his great support towards the success of this work.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Usiabulu G Idanegbe, Azubuike H Amadi, Emeka J Okafor, Jumbo-Egwurugwu Precious. Optimization of methane and natural gas liquid recovery in a reboiled absorption column International Journal of Scientific Research and Engineering Development. IJSRED. 2022;5(2):36-48.
2. Tan SY, Tan S. The Right Technologies for Gas Leak Detection. 2019;43.
3. Nosike L. A study of gas fire at a water well in the Caritas University Premises, Amorji Nike, Enugu, PTD Journal Review; 2023.
4. Appah D, Aimikhe V, Okologume W. Assessment of Gas Leak Detection Techniques in Natural Gas Infrastructure. Retrieved from OnePetro; 2021. Available:https://doi.org/10.2118/208236-MS
5. Teriba A. Harmonization of Fiscal and Monetary Policies in Nigeria; 2018. Available:https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3163481
6. The Gas Index. (n.d.). Retrieved from Where Leaks Occur. Available:https://thegasindex.org/where-leaks-occur/
7. Kedar Potdar, Rishab Kinnerkar. A Comparative Study of Machine Learning Algorithms applied to Predictive Breast Cancer Data; 2015.
8. Nosike L. Exploration and production Geoscience-Comprehensive Skills Acquistion for an Evolving industry, Delizon Publishers. 2020;441:227-264.
9. Nosike L. Relationship between tectonics and vertical hydrocarbon leakage: PhD thesis, University of Nice-Sophia Antipolis, Fracnce. 2009;281.
10. Godsday Idanegbe Usiabulu, Azubuike Hope Amadi, Oluwatayo Adebisi, Uchenna Donald Ifedili, Kehinde Elijah Ajayi, Pwafureino Reuel Moses. Gas flaring and its environmental impact in Ekpan Community, Delta state, Nigeria, American Journal of Science Engineering and Technology. 2023;8(1):42-53.
11. Godsday Idanegbe Usiabulu, Azubuike Hope Amadi, Emeka Okafor, Jumbo-Egwurugwu Preciuos, Optimization of methane and natural gas liquid recovery in a reboiled absorbtion column. International Journal of Scientific Research and Engineering Development. 2022;5(2).
12. Chinwuko Emmanuel Chuka, Ifowodo Henry Freedom, Umeozokwere Anthony O. Transient Model-Based Leak Detection and Localization Technique for Crude Oil Pipelines: A Case of N.P.D.C, Olomoro; 2016.
13. Kedar Potdar, Rishab Kinnerkar. A Comparative Study of Machine Learning Algorithms applied to Predictive Breast Cancer Data; 2013.
14. Zukang Hu, Beiqing Chen, Wenlong Chen, Debao Tan and Dingtao Shen. Review of model-based and data-driven approaches for leak detection and location in water distribution systems; 2021.

SUPPLEMENTARY MATERIALS

Coding for Machine Learning and Automation

```
In [7]: print('min Time value', df.Time.min())
        print('max Time value', df.Time.max())

        min Time value 1000
        max Time value 9000

In [8]: #### Select relevant columns
        relevant_df = df[['Time','Pr_final','Pr_initial','Tolerance','Min','Max']]
        cleaned_df = relevant_df.dropna()
        print(relevant_df.shape)
        print(cleaned_df.shape)
        print('{} rows dropped from the table'.format(relevant_df.shape[0]-cleaned_df.shape[0]))

        (1012, 6)
        (1001, 6)
        11 rows dropped from the table

In [9]: cleaned_df['Tolerance'] = cleaned_df['Tolerance'].astype(float)
```

Plotting

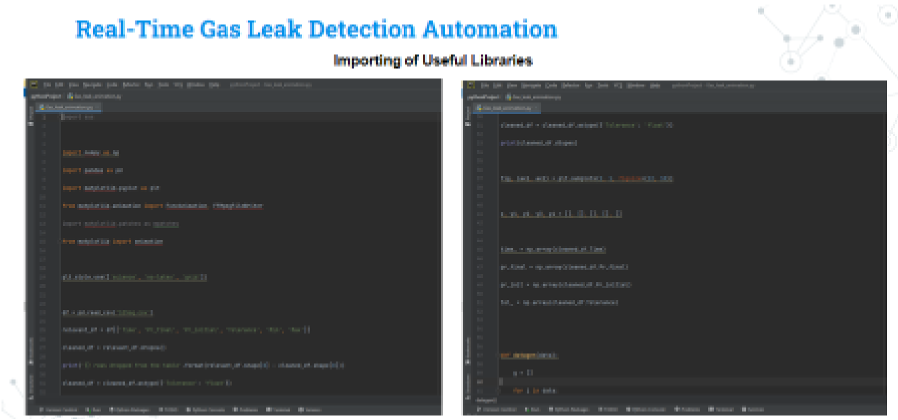
```
In [10]: # adding some nice colors
         plt.rcParams['text.color'] = 'black'
         plt.rcParams['axes.labelcolor'] = 'blue'
         plt.rcParams['xtick.color'] = 'red'
         plt.rcParams['ytick.color'] = 'red'


         fig, axs = plt.subplots(2, figsize=(12,10))
         fig.tight_layout(pad=1.08, h_pad=7, w_pad=None)

         axs[0].plot(cleaned_df.Time, cleaned_df.Pr_final,'-p',
                     lw=1.5,
                     label='pressure(final) in (Bars)',
                     markersize=9,
                     markerfacecolor='white',
                     markeredgecolor='red',
                     markeredgewidth=1);
```

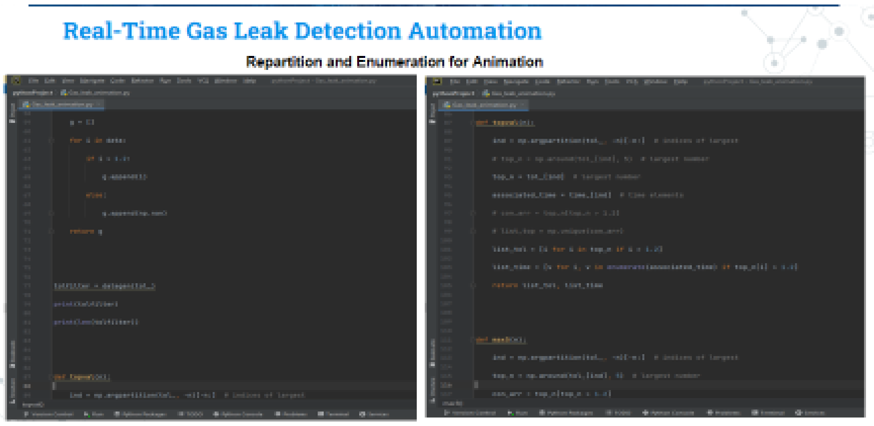
```
In [1]: import warnings
        warnings.filterwarnings("ignore")

In [2]: import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
        from pylab import cm
        from matplotlib.ticker import MaxNLocator
        from matplotlib.ticker import FormatStrFormatter

In [3]: #pip install git+https://github.com/garrettj403/SciencePlots

In [4]: plt.style.use(['science','no-latex','grid'])

In [5]: df = pd.read_csv('iESog.csv')

In [6]: df.head()
```

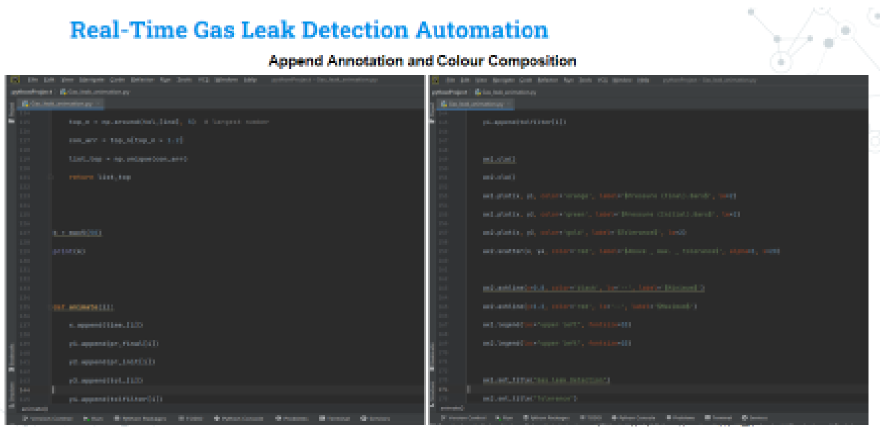
Out[6]:

	Time	Pr_final	Pr_initial	Events	Tolerance	Min	Max
0	4005	3.5	1.5	Residual stage	0.428571429	0.8	1.2
1	4010	3.5	1.5	NaN	0.428571429	0.8	1.2
2	4015	3.5	1.5	NaN	0.428571429	0.8	1.2
3	4020	3.5	1.5	NaN	0.428571429	0.8	1.2
4	4025	3.5	1.5	NaN	0.428571429	0.8	1.2

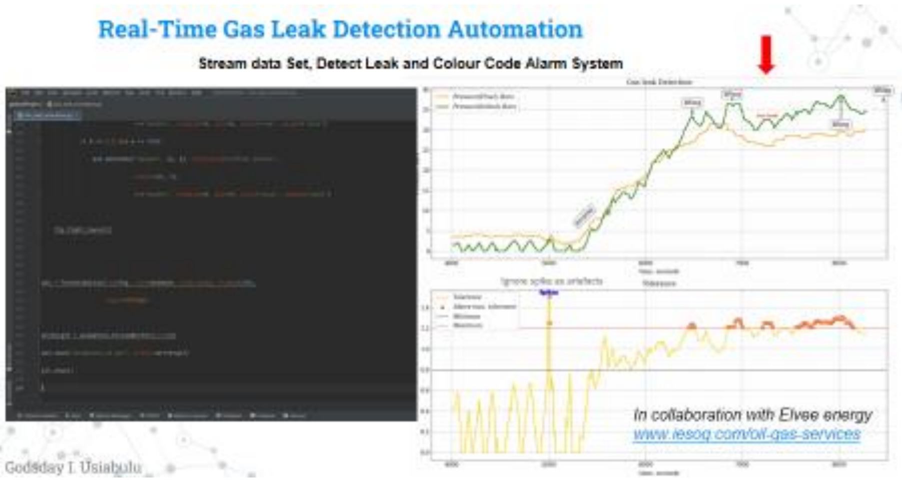
```
In [7]: print('min Time value', df.Time.min())
        print('max Time value', df.Time.max())

        min Time value 1000
        max Time value 9000

In [8]: #### Select relevant columns
        relevant_df = df[['Time','Pr_final','Pr_initial','Tolerance','Min','Max']]
        cleaned_df = relevant_df.dropna()
```


Real-Time Gas Leak Detection Automation
Importing of Useful Libraries


Real-Time Gas Leak Detection Automation
Repartition and Enumeration for Animation


Real-Time Gas Leak Detection Automation
Append Annotation and Colour Composition


Real-Time Gas Leak Detection Automation
Append Annotation and Separate Leak from Lifting



Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/102508