Indian Journal of Petroleum Engineering (IJPE)
ISSN: 2582-9297 (Online), Volume-6 Issue-1, May 2026

A Machine Learning–Based Virtual Flow Meter for Continuous Estimation of Well Production Rates

Mohammad Amir Ashraff

Abstract: ***Accurate estimation of liquid rate, gas rate, and water cut is essential for effective production surveillance, reservoir management, and operational decision-making in oil and gas assets. In most producing fields, direct production measurements are obtained through periodic well tests or selectively deployed multiphase flow meters, resulting in sparse temporal resolution, delayed detection of production changes, and limited field-wide visibility. Physics-based production models provide valuable engineering insight but require frequent calibration and often struggle to maintain accuracy under transient operating conditions and evolving reservoir behavior. These limitations motivate the use of data-driven approaches that leverage existing field instrumentation to deliver continuous production estimates. This paper presents a machine learning–based Virtual Flow Meter (VFM) for continuous estimation of oil, gas, and water production rates using routinely available operational measurements. High-frequency field data, including pressures, temperatures, choke settings, and, where applicable, lift-gas injection rates, are temporally aligned with historical well-test and laboratory measurements to construct reliable training datasets. Independent regression models are developed for oil, gas, and water rates, allowing each model to capture phase-specific sensitivities while maintaining physical consistency and avoiding reliance on explicit flow-regime identification or mechanistic multiphase correlations. The proposed VFM is deployed end-to-end on a commercial data analytics platform, enabling continuous ingestion of sensor data, real-time inference, performance monitoring, and periodic retraining. Model validation using historical field data demonstrates strong agreement between predicted and reference production values across all phases, indicating that the data- driven VFM provides reliable, meter-like production estimates. The results show that continuous production surveillance can be achieved without additional hardware or instrumentation, offering a scalable and cost-effective solution for large well portfolios. This approach supports proactive operational decision- making and enhances production visibility across modern oil and gas assets.***

Index Terms: ***Virtual Flow Meter, Machine Learning, Well Rate Estimation, Multiphase Flow, Production Surveillance.***

Nomenclature:
VFM: Virtual Flow Meter
MPFM: Multiphase Flow Meter
ML: Machine Learning
BHP: Bottom Hole Pressure
WC: Water Cut
IoT: Internet of Things

Manuscript received on 09 January 2026 | First Revised Manuscript received on 19 January 2026 | Second Revised Manuscript received on 16 April 2026 | Manuscript Accepted on 15 May 2026 | Manuscript published on 30 May 2026.
* Correspondence Author(s)

Mohammad Amir Ashraff*, Department of Data Science and Advanced Analytics, Hyderabad (Telangana), India. Email ID: amir.ashraff@bdb.ai, ORCID ID: 0009-0006-9735-1190

I. INTRODUCTION

Production rate estimation underpins almost every engineering and operational decision across the lifecycle of an oil and gas asset. Accurate knowledge of liquid rate, gas rate, and water cut is required for reservoir surveillance, production optimisation, artificial lift tuning, and surface facility management. Despite its importance, continuous and reliable production measurement remains challenging due to the multiphase nature of produced fluids and the distributed architecture of production systems.

Traditional production measurement approaches rely on periodic well tests or physical metering devices. These methods provide accurate point measurements but lack continuous visibility and are challenging to scale across large well populations. As digital oilfield initiatives mature, there is growing interest in data-driven alternatives capable of inferring production behaviour from existing instrumentation, as reviewed in prior VFM literature [1].

From an operational standpoint, the limitation of traditional approaches is not accuracy in isolation, but the inability to provide continuous, asset-wide production visibility at scale. In modern brownfield assets, where instrumentation density has increased significantly, large volumes of high-frequency pressure, temperature, and flow-related measurements are already available but remain underutilised for production estimation. Data-driven Virtual Flow Meters have been proposed to exploit such latent operational information, enabling a shift from sparse, test-based production assessment to continuous, real-time production intelligence.

This work presents a fully integrated, end-to-end data-driven Virtual Flow Meter that simultaneously estimates oil, gas, and water rates and is deployed in a production environment. The solution enables continuous ingestion of operational sensor data, automated model training and inference, performance monitoring, retraining, and real-time production surveillance.

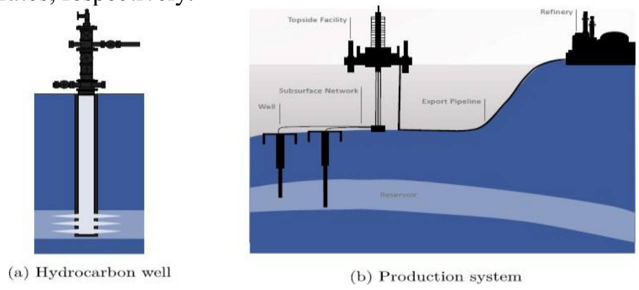
II. PHYSICAL PRODUCTION SYSTEMS

Oil and gas production systems transport hydrocarbons from subsurface reservoirs to surface processing facilities under complex thermodynamic and hydraulic conditions. The produced stream is inherently multiphase, typically consisting of oil, gas and water flowing simultaneously through the reservoir, wellbore and surface network.

At any location, the total volumetric production rate is given by

$$Q_{\text{total}} = Q_O + Q_g + Q_W, \quad \dots \quad (1)$$

where Q_o, Q_g and Q_w denote oil, gas and water flow rates, respectively.



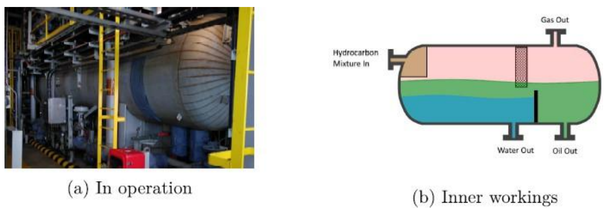
[Fig.1: End-to-End Hydrocarbon Production System Illustrating Multiphase Flow from Reservoir to Surface Facilities (Adapted from [2])]

III. SURFACE PROCESSING AND PHASE SEPARATION

Once produced fluids reach the surface, they must be separated into gas, oil and water streams. Water cut, a key production indicator, is defined as

$$WC = \frac{Q_w}{Q_o + Q_w} \quad \dots \quad (2)$$

Horizontal three-phase separators perform phase separation based on gravity settling and residence time. However, transient flow behaviour and varying inlet conditions introduce uncertainty into the measured phase rates.



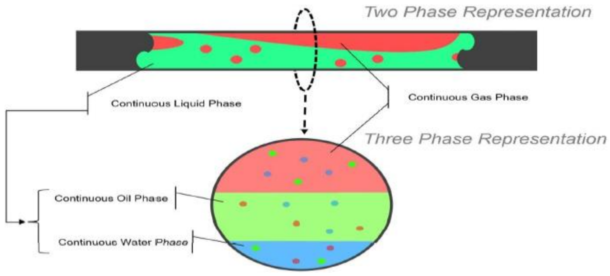
[Fig.2: Horizontal Three-Phase Separator Illustrating Gravity-Based Phase Separation (Adapted from [2])

IV. MULTIPHASE FLOW REGIME VARIABILITY

Multiphase flow exhibits a wide range of flow regimes as a function of phase fractions, velocities, and pipe inclination. These regimes affect pressure losses and phase distribution, leading to a non-unique relationship between measured pressures and actual flow rates.

The pressure gradient along the flow path is expressed as

$$\frac{dP}{dz} = \frac{dP}{dz}_{fric} + \frac{dP}{dz}_{grav} + \frac{dP}{dz}_{acc} \quad \dots \quad (3)$$



[Fig.3: Representative Multiphase Flow Regimes in Horizontal Pipelines (Adapted from [2])]

V. TEMPORAL VARIABILITY AND SPARSE GROUND TRUTH

Production rates evolve continuously due to reservoir depletion and operational changes:

$$\frac{dQ(t)}{dt} <> 0. \quad \dots \quad (4)$$

In contrast, direct production measurements are available only at discrete well-test times:

$$Q_{\text{meas}}(t) = Q(t_k) \quad \dots \quad (5)$$

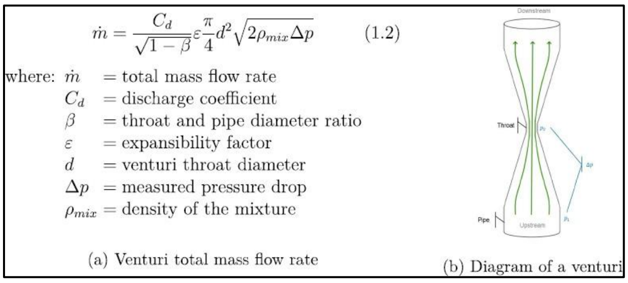
[Fig.4: Petroleum Lifecycle Illustrating Persistent Need for Production Rate Estimation (Adapted from [2])]

VI. TRADITIONAL PRODUCTION RATE ESTIMATION METHODS

A. Multiphase Flow Metering

Venturi-based multiphase flow meters estimate total mass flow as

$$m = CA\sqrt{2\rho \triangle P} \quad \dots \quad (6)$$



[Fig.5: Venturi-Based Multiphase Flow Metering Principle (Adapted from [2])]

In practice, the accuracy of multiphase flow meters is strongly influenced by flow regime, phase slip and fluid property variability. Most MPFM technologies rely on combinations of differential pressure measurements, gamma-ray attenuation, impedance or microwave-based sensors [1], coupled with slip models to separate phase contributions. These models require calibration against reference well tests and are sensitive to changes in gas volume fraction, water cut, and flow-regime transitions. Under transient operating conditions, such as choke adjustments or slug flow, the performance of hybrid and model-based flow estimation approaches can degrade due to sensitivities in calibration and modelling. As a result, while MPFMs provide reliable localised measurements when properly calibrated, their accuracy and transferability across wells and over time remain challenging, limiting their scalability as a field-wide continuous production monitoring solution. Data-driven Virtual Flow Meters address these limitations by learning production behaviour directly from historical operational data, thereby reducing dependence on

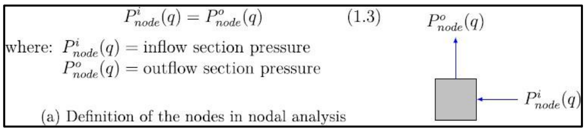
Indian Journal of Petroleum Engineering (IJPE)
ISSN: 2582-9297 (Online), Volume-6 Issue-1, May 2026

explicit flow regime identification and slip modelling.

B. Physics-Based Production Modelling

Physics-based production models estimate flow rates by solving pressure-balance equations. In nodal analysis, the operating point satisfies.

$$P_r - \Delta P_{in}(Q) - \Delta P_{out}(Q) = P_{s.} \quad \dots \quad (7)$$



[Fig.6: Physics-Based Nodal Analysis Framework for Production Rate Estimation (Adapted From [2])]

Physics-based production models estimate flow rates by coupling inflow performance relationships with wellbore and surface pressure-loss correlations. These models rely on empirical or semi-mechanistic correlations to account for multiphase pressure drop and flow-regime effects. While highly effective for design studies, sensitivity analysis and scenario planning, such models require frequent recalibration to remain accurate under changing operating conditions. Variations in water cut, gas–liquid ratio, artificial lift settings, and reservoir depletion can introduce significant mismatches between modelled and observed behaviour. Consequently, physics-based models are best suited for engineering analysis and planning, rather than continuous, real-time production surveillance across large well populations. Data-driven Virtual Flow Meters complement these models by providing adaptive, continuously updated production estimates without requiring repeated manual recalibration.

VII. LIMITATIONS OF CONVENTIONAL APPROACHES

Conventional approaches are constrained by high cost, sparse temporal resolution, calibration effort and limited scalability. These limitations motivate data-driven alternatives that leverage existing instrumentation to continuously infer production behaviour.

In practice, physical multiphase flow meters are typically deployed at a limited number of locations due to their high capital cost, complex installation, and ongoing maintenance requirements. As a result, most wells in a field continue to rely on infrequent well tests for production estimation, leaving long periods of operational uncertainty. Physics-based production models, while valuable for design and planning, depend heavily on assumptions about flow regimes, fluid properties, and boundary conditions that may not hold during transient operations, such as choke adjustments or artificial lift optimisation.

These limitations create a mismatch between the need for continuous, real-time production awareness and the capabilities of conventional approaches. Data-driven Virtual Flow Meters address this gap by providing continuous estimates across all wells using existing sensor infrastructure, thereby extending production surveillance beyond the physical reach of metering hardware. From an asset economics perspective, the value of a production measurement solution is determined not only by per-well accuracy but also by its ability to deliver continuous visibility across the entire healthy portfolio at a sustainable cost. Physical multiphase flow meters typically entail high capital expenditure, specialised installation, and recurring maintenance, which limit their deployment to a small subset of strategic wells. Physics-based models, while less capital-intensive, require significant engineering effort for calibration and ongoing maintenance as operating conditions evolve.

In contrast, data-driven Virtual Flow Meters leverage existing sensor infrastructure, resulting in near-zero incremental hardware cost and minimal marginal cost per additional well. Once deployed, the same modelling framework can be scaled across hundreds of wells with consistent performance. This shift from asset-level instrumentation to portfolio-level analytics fundamentally changes the economics of production surveillance, enabling operators to achieve continuous, field-wide rate estimation that would be economically impractical with conventional metering approaches.

VIII. DATA-DRIVEN VIRTUAL FLOW METER METHODOLOGY

A data-driven Virtual Flow Meter (VFM) estimates well production rates by learning the relationship between routinely measured operational parameters and reference production data, without relying on explicit physical flow models or dedicated metering hardware. Unlike conventional approaches that model multiphase flow behaviour through analytical correlations or mechanistic assumptions, a data-driven VFM treats the production system as a black box mapping observable inputs to production outputs. This paradigm has been widely discussed in the literature as suitable for oil and gas production systems, where multiphase flow behaviour is highly nonlinear, regime-dependent, and sensitive to operational disturbances [1].

Let $\mathbf{x}_{w,t}$ denote the vector of operational measurements for well w at time t, including pressures, temperatures, choke settings and lift-gas injection rates where applicable. The objective of the VFM is to estimate the corresponding production rates $\mathbf{y}_{w,t}$ by learning a nonlinear functional relationship:

$$Y_{w,t} = F\left(X_{w,t}\right), \quad \dots \quad (8)$$

where $F(\cdot)$ is a data-driven regression model trained using historical data. In this work, the output vector is defined as

$$\hat{\mathbf{y}}_{w,t} = \begin{bmatrix} \hat{Q}_o \\ \hat{Q}_g \\ \hat{Q}_w \end{bmatrix}, \quad \dots \quad (9)$$

representing estimated oil, gas and water production rates, respectively.

A. Rationale for a Data-Driven Formulation

The motivation for adopting a data-driven formulation stems from the inherent limitations of physics-based and hybrid models in operational settings. Physics-based VFMs require explicit identification of flow regimes, accurate characterisation of fluid properties, and careful calibration of pressure-loss correlations [1]. In practice, these requirements are challenging to satisfy consistently across large well populations and over extended periods of operation. Slight deviations in fluid composition, sensor drift, or operational conditions can lead to significant estimation errors.

In contrast, data-driven VFMs implicitly capture the combined effects of flow-regime transitions, fluid-property variations, and operational constraints directly from historical data. As long as the training dataset adequately spans the operating envelope of the well, the learned mapping remains valid even under transient conditions such as choke changes, lift-gas adjustments, or gradual reservoir depletion. This makes data-driven VFMs particularly attractive for mature brownfield assets, where large volumes of historical data are available but physical models struggle to remain well-calibrated under changing operating conditions.

B. Training Data Construction and Alignment

Operational measurements in the field are typically recorded at high frequency, often at intervals of five minutes or less. Reference production data, however, are obtained from well tests or laboratory analyses and are available only at discrete points in time. To reconcile this mismatch, operational data are aggregated into daily averages and strictly aligned with the dates for which reliable reference measurements exist.

Only data records that satisfy predefined quality criteria—such as sensor availability, stable operating conditions, and validated test results—are retained for model training. This filtering step is critical to prevent the model from learning noise and measurement artefacts. By enforcing strict data integrity checks, the resulting training dataset provides a consistent mapping between operational states and actual production outcomes.

C. Model Decomposition and Output Strategy

Rather than training a single multi-output model, the proposed VFM architecture employs three independent regression models [3]:

$$Q_o = f_o(x),\ Q_g(X),\ \ Q_g = f_g(X), Q_w = f_w(X).\ \ldots \quad (10)$$

This decomposition allows each model to specialise in learning the sensitivities relevant to a specific production phase. Independent phase-specific models have been reported to improve learning flexibility in prior machine-learning-based flow-estimation studies compared to a single coupled formulation.

Each model is trained by minimizing a mean squared error objective:

$$\mathcal{L}_k \frac{1}{N}\sum_{i=1}^{N}\left(y_i{}^{(k)} - \hat{y}_i{}^{(k)}\right)^2, k \epsilon \{o, g, w\} \ \ldots \quad (11)$$

During training, additional physical consistency constraints are enforced to ensure that predicted rates are nonnegative and water-cut estimates are valid.

D. Generalization and Operational Robustness

A key requirement for operational deployment is the VFM's ability to generalise beyond the exact conditions observed during training. Data-driven VFMs achieve this by learning smooth nonlinear mappings rather than brittle rule-based relationships. When trained on sufficiently diverse historical data, the model reliably interpolates across unseen combinations of pressure, temperature, and choke settings.

Furthermore, because the VFM does not depend on explicit flow regime classification, it remains stable across regime transitions that often cause discontinuities in physics-based models [1]. This property is significant for wells operating near regime boundaries or under variable artificial lift conditions.

E. Inference and Deployment Considerations

Once trained, the VFM operates in inference mode by consuming real-time operational measurements and producing continuous production estimates at the exact temporal resolution of the input data. The computational cost of inference is negligible compared to physical simulation, enabling deployment at scale across hundreds or thousands of wells.

In practice, the VFM outputs are integrated into production surveillance dashboards, where engineers can monitor trends in oil, gas, and water production alongside pressure and temperature measurements. This tight integration enables rapid anomaly detection, improved well-performance diagnostics, and more-informed operational decision-making.

F. When Not to Use a Data-Driven Virtual Flow Meter

Despite their advantages, data-driven Virtual Flow Meters are not universally applicable and should not be viewed as a replacement for all production measurement technologies. In early-life fields or greenfield developments, where historical production data are sparse or nonexistent, physics-based models remain essential for initial production forecasting and facility design. Similarly, in scenarios requiring custody transfer or regulatory-grade measurement, physical flow meters remain the industry standard due to their traceability and compliance requirements.

Data-driven VFMs also depend critically on the quality and representativeness of the training dataset. Wells operating far outside the historical operating envelope or subject to significant hardware changes, such as recompletions or artificial lift modifications, may require model retraining or a temporary fallback to conventional estimation methods. For these reasons, data-driven VFMs are best positioned as a complementary technology—optimally suited for continuous production

Retrieval Number: 100.1/ijpe.A192406010526
DOI: 10.54105/ijpe.A1924.06010526
Journal Website: www.ijpe.latticescipub.com

Published By:
Lattice Science Publication (LSP)

Indian Journal of Petroleum Engineering (IJPE)
ISSN: 2582-9297 (Online), Volume-6 Issue-1, May 2026

surveillance and operational decision support—rather than as a universal replacement for physics-based or metering solutions.

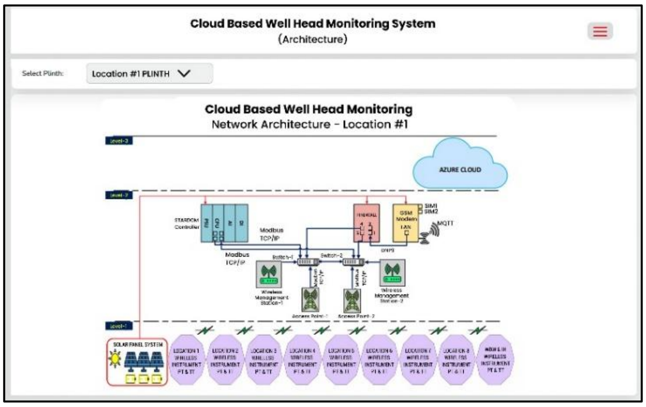
G. Platform Architecture and Deployment

The proposed data-driven Virtual Flow Meter has been implemented and deployed end-to-end on a commercial data analytics platform, enabling continuous production monitoring from raw field data ingestion to decision-ready visualisation. The platform provides an integrated environment for data ingestion, feature engineering, model building, training, inference, monitoring, and lifecycle management, which is critical for operationalising machine learning solutions in production oil and gas environments.

High-frequency operational measurements from field sensors are streamed into the platform via industrial IoT pipelines, ensuring reliable, low-latency ingestion of pressure, temperature, choke, and lift-gas signals. These measurements are automatically validated, aggregated and transformed into model-ready features using standardized data pipelines. The trained Virtual Flow Meter models consume these features in near real time to generate continuous estimates of oil, gas, and water rates for each well.

Beyond inference, the platform supports end-to-end governance of models. Prediction outputs are continuously tracked against available reference measurements, enabling automated performance monitoring and early detection of model drift or abnormal healthy behaviour. Retraining workflows can be triggered periodically or conditionally based on performance thresholds, ensuring sustained accuracy as reservoir conditions and operating regimes evolve.

Model outputs are surfaced through interactive production surveillance dashboards that provide engineers and operations teams with intuitive access to trends, anomalies and well-level performance indicators. This tight integration between data ingestion, machine learning, and visualisation allows the Virtual Flow Meter to function not as an offline analytical tool, but as a continuously operating digital production measurement system deployed at scale across the asset.



[Fig.7: End-to-End Virtual Flow Meter Architecture on a Commercial Data Analytics Platform]

The following section evaluates the predictive performance of this deployed system using historical field data.

IX. OVERALL ARCHITECTURE

The diagram represents a three-level industrial IoT / SCADA architecture for remote wellhead monitoring, where:

- Field instruments collect pressure and temperature
- Controllers and network devices aggregate and secure data
- Cloud (Azure) receives data for storage, analytics, dashboards, and Virtual Flow Meter (VFM) models

A. LEVEL-1 - Field / Instrumentation Layer

i. *Wireless Instrument (PT & TT)*

- PT = Pressure Transmitter
- TT = Temperature Transmitter
- These are field sensors mounted at each wellhead location.
- They measure:
- Wellhead pressure
- Wellhead temperature
- Each location (Location 1 to 8, WDW (Wellhead Data Wireless Unit) and IH (Injection Head)) has its own wireless instrument.

ii. *Wireless*

- Indicates wireless communication from instruments to access points.
- Typically, ISM (Industrial, Scientific and Medical - originally allocated for equipment such as microwave ovens, MRI machines, and industrial heaters; today widely used for wireless data communication) band radios operating at 900MHz, 2.4GHz, or proprietary industrial frequencies.

iii. *Solar Panel System*

- Powers the wireless instruments and field electronics.
- Includes:
- Solar panels
- Battery bank
- Charge controller
- Used because wellheads are often located in remote areas without grid power.

iv. *Lightning Symbol*

- Represents electrical power or energy flow, or field signal activity.
- Commonly used in industrial drawings to indicate powered devices.

B. LEVEL-2 - Control, Networking & Edge Layer

This layer forms the backbone for industrial control and communication.

i. *STARDOM Controller*

A Remote Terminal Unit (RTU) / PLC (Programmable Logic Controller) class controller.
Internal modules shown:

- PSU - Power Supply Unit
- CPU - Central Processing Unit

- AI - Analogue Inputs (sensor data input)
- DI - Digital Inputs

ii. *Function:*

- Collects data from wireless instruments
- Executes basic logic
- Publishes data to higher systems

iii. *Modbus TCP/IP*

- Industrial communication protocol
- Used between:
- RTU and switches
- RTU and access points
- Runs over Ethernet using TCP/IP

iv. *Switch-1 / Switch-2*

- Industrial Ethernet switches
- Aggregate data from:
- Access points
- Controllers
- Provide network segmentation and redundancy

v. *Wireless Management Station-1 / Wireless Management Station-2*

- Used to:
- Configure wireless radios
- Monitor signal health
- Manage radio networks
- Often vendor-specific tools

vi. *Access Point-1 / Access Point-2*

- Industrial wireless access points
- Act as gateways between:
- Field wireless instruments
- Wired Ethernet network
- Provide redundancy and coverage

vii. *Firewall*

- Cybersecurity device
- Protects the OT (Operational Technology) network
- Controls:
- Inbound traffic
- Outbound traffic
- Prevents unauthorized access to wellhead systems

viii. *GSM Modem*

- GSM = Global System for Mobile Communications
- Cellular modem used for:
- Remote connectivity
- Internet backhauls
- Uses mobile networks instead of leased lines

ix. *SIM1 / SIM2*

- Dual SIM cards
- Used for:
- Network redundancy
- Failover between telecom operators
- Ensures higher uptime

x. *LAN*

- LAN = Local Area Network
- Internal Ethernet network connecting:
- Controllers
- Switches
- Firewall
- GSM modem

xi. *DNP3*

- Distributed Network Protocol v3
- Widely used SCADA protocol
- Often used between:
- RTUs
- Control systems
- Supports secure telemetry and control

xii. *MQTT*

- Message Queuing Telemetry Transport
- Lightweight IoT protocol
- Used for:
- Sending data to the cloud
- Low-bandwidth, reliable messaging
- Ideal for remote oilfield deployments

C. LEVEL-3 — Cloud / Enterprise Layer

i. *Azure Cloud*

- Microsoft Azure Cloud Platform
- Receives data via MQTT
- Used for:
- Data storage
- Dashboards
- Analytics
- Machine Learning
- Virtual Flow Meter (VFM) models
- Enables remote monitoring and decision-making

ii. *Additional Labels & Symbols/ Location #1 PLINTH*

- Physical installation base for equipment
- Usually, a concrete foundation holding:
- RTU
- Switches
- Power systems

iii. *Red Boundary Box*

- Represents secured OT (Operational Technology) network zone
- Everything inside is protected behind firewall rules

Operational Technology (OT) refers to the hardware and software systems used to monitor, control, and operate physical industrial processes in real time. In oil and gas, OT includes:

- Wellhead instruments (PT, TT, flow sensors)
- PLCs / RTUs
- Controllers (e.g., wellhead controllers)
- SCADA systems
- Industrial networks (Modbus, DNP3, OPC)

iv. *Dashed Horizontal Lines*

- Clearly separate architecture layers:
- Level-1 - Field
- Level-2 - Control / Edge
- Level-3 - Cloud

v. *Virtual Flow Meter Perspective*

- Inputs
- PT, TT, choke position, operating conditions
- Transport:
- Wireless → RTU → MQTT → Azure
- Processing:

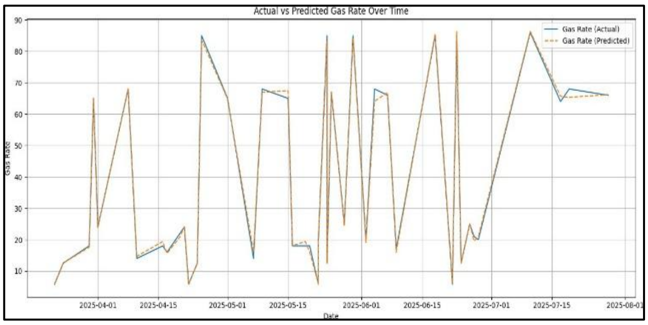
Retrieval Number: 100.1/ijpe.A192406010526
DOI: 10.54105/ijpe.A1924.06010526
Journal Website: www.ijpe.latticescipub.com

Published By:
Lattice Science Publication (LSP)

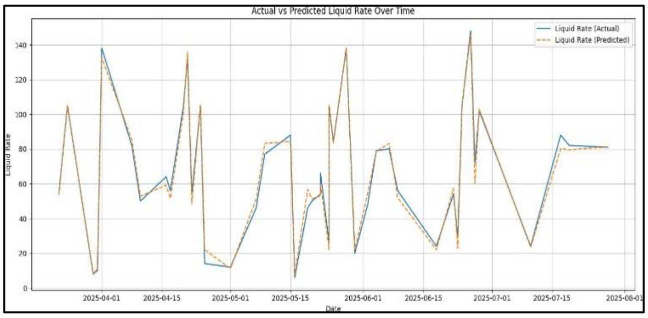
- Cloud analytics and machine learning models
- Outputs:
- Continuous oil, gas, and water production estimates
- Dashboards and alerts

The following section evaluates the predictive performance of this deployed system using historical field data.

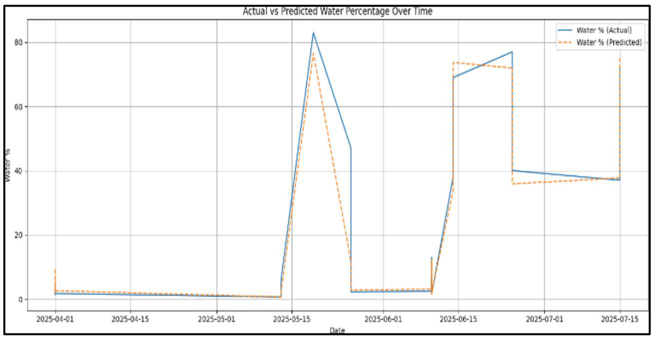
X. MODEL VALIDATION USING HISTORICAL DATA



[Fig.8: Actual Versus Predicted Gas Rate Over Time]



[Fig.9: Actual Versus Predicted Liquid Rate Over Time]



[Fig.10: Actual Versus Predicted Water Percentage Over Time]

Similar levels of agreement between predicted and reference rates have been reported in prior machine-learning-based studies of multiphase flow estimation.

XI. POSITIONING WITHIN EXISTING LITERATURE

Virtual Flow Meters can be broadly categorised as physics-based, hybrid, or data-driven. Recent literature emphasises that purely data-driven VFMs offer superior adaptability and scalability in complex production environments [1], where multiphase flow behaviour is highly nonlinear and complex to model explicitly.

While physics-based and hybrid VFMs remain indispensable for design studies, flow assurance analysis and regulatory reporting, data-driven VFMs are increasingly recognized as the preferred solution for continuous production surveillance and operational decision support. In assets with sufficient historical data and sensor coverage, data-driven approaches offer a favourable balance between accuracy, coverage and lifecycle cost.

XII. DISCUSSION

A. Industrial Deployment Considerations

Production surveillance in most operating assets remains constrained by sparse well tests and limited deployment of physical multiphase flow meters, resulting in delayed visibility into well performance and production anomalies. The Virtual Flow Meter presented in this work addresses this gap by enabling continuous, field-wide estimation of oil, gas and water production rates using existing sensor infrastructure.

Implemented and deployed on a commercial data analytics platform, the solution delivers an end-to-end operational capability spanning continuous ingestion of IoT sensor data, automated model building, training, inference, performance monitoring, retraining, and real-time production dashboards. This platform-enabled deployment transforms the VFM from a standalone modelling exercise into a scalable production monitoring system, allowing operators to move from reactive, test-driven workflows to proactive, data-driven production management across large well portfolios.

The proposed VFM provides continuous production visibility using existing instrumentation and demonstrates strong agreement with reference measurements. Its scalability and low deployment cost make it particularly attractive for mature fields with large well populations. Limitations related to data quality and concept drift can be mitigated through periodic retraining and performance monitoring.

From a decision-making perspective, the Virtual Flow Meter enables operators to move from reactive, test-driven workflows to proactive production management. Continuous rate estimates support faster detection of production anomalies, improved choke and lift optimization and more informed allocation decisions across wells. These operational benefits are difficult to achieve with sparse well tests or with physical meters deployed selectively alone.

Model performance is continuously monitored in production, with scheduled retraining triggered by data drift or degradation alerts, in line with established best practices for industrial machine learning systems.

Viewed together, physical multiphase flow meters provide localized measurement capability, physics-based models offer mechanistic insight for design and analysis, and data-driven Virtual Flow Meters bridge the gap by enabling scalable, continuously adaptive production surveillance across entire

Retrieval Number: 100.1/ijpe.A192406010526
DOI: 10.54105/ijpe.A1924.06010526
Journal Website: www.ijpe.latticescipub.com

Published By:
Lattice Science Publication (LSP)

well portfolios.

B. End-to-End Operational Deployment

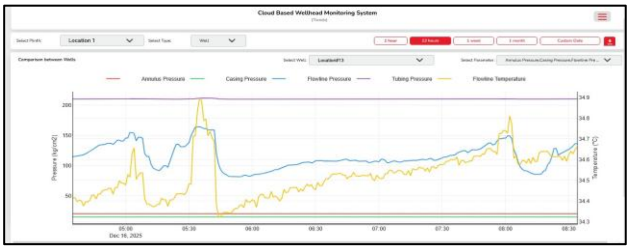
Beyond model development, a critical contribution of this work is the successful end-to-end operationalization of the Virtual Flow Meter on a commercial data analytics platform. The platform provides an integrated environment for data ingestion, model deployment, monitoring and lifecycle management, enabling the proposed VFM to function as a continuously operating production surveillance solution rather than an offline analytical model.

Field instrumentation across wells streams high-frequency pressure, temperature and operational data through industrial IoT pipelines into a commercial data analytics platform. These raw sensor feeds are automatically validated, aggregated and transformed into model-ready features. The trained VFM models are deployed as production services on the platform, generating continuous estimates of oil, gas, and water rates in near-real time.

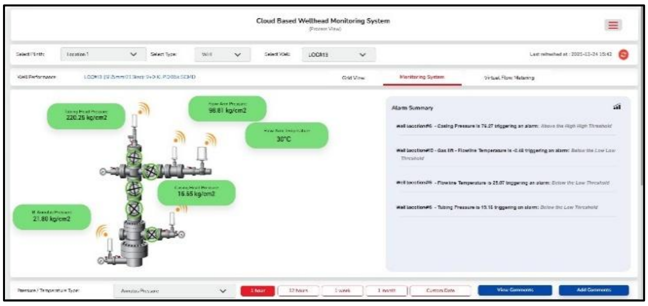
Operational dashboards built on the platform provide engineers with transparent visibility into predicted production trends alongside key input variables. Model performance is continuously monitored against reference data, where available, with automated alerts triggered for abnormal deviations or potential data quality issues. This closed-loop monitoring framework ensures early detection of model drift, sensor anomalies, or changes in healthy behaviour.

Model retraining and version management are handled natively within the platform, allowing updated models to be validated and redeployed with minimal manual intervention. As a result, the complete VFM lifecycle—from raw sensor data ingestion to production dashboards and alerts—is managed within a single, unified system. This end-to-end capability significantly lowers the adoption barrier and distinguishes the proposed solution from prior VFM studies that focus primarily on model development, without addressing operational deployment.

[Fig.11: Production Surveillance Dashboard on a Commercial Data Analytics Platform Showing Real-Time VFM Predictions]



[Fig.12: Production Surveillance Dashboard on a Commercial Data Analytics Platform Showing Input Sensor Trends]



[Fig.13: Production Surveillance Dashboard on a Commercial Data Analytics Platform Showing Model Performance Monitoring]

XIII. CONCLUSION

This paper presented a machine-learning–based Virtual Flow Meter for continuous estimation of liquid rate, gas rate, and water percentage. By leveraging existing operational measurements and historical production data, the proposed approach overcomes the limitations of conventional measurement techniques. It provides a scalable, cost-effective solution for modern oil and gas assets.

The results indicate that, for production surveillance and operational decision support, data-driven Virtual Flow Meters can replace physics-based models and complement physical multiphase flow meters in the majority of field scenarios. As digital transformation initiatives continue to expand across the oil and gas industry, data-driven VFMs are well-positioned to become a core component of future production-monitoring architectures.

Implemented on a commercial data analytics platform, the proposed Virtual Flow Meter operates as a fully managed, end-to-end production intelligence solution—spanning continuous sensor ingestion, real-time prediction, performance monitoring, automated alerting, and periodic retraining.

DECLARATION STATEMENT

I must verify the accuracy of the following information as the article's author.

- **Conflicts of Interest/ Competing Interests:** Based on my understanding, this article has no conflicts of interest.
- **Funding Support:** This article has not been funded by any organizations or agencies. This independence ensures that the research is conducted objectively and without external influence.
- **Ethical Approval and Consent to Participate:** The content of this article does not necessitate ethical approval or consent to participate with supporting documentation.
- **Data Access Statement and Material Availability:** The adequate resources of this article are publicly accessible.
- **Author's Contributions:** The

authorship of this article is contributed solely.

REFERENCES

1. T. Bikmukhametov and J. Ja¨schke, "First principles and machine learning virtual flow metering: A literature review," *Journal of Petroleum Science and Engineering*, vol. 184, p. 106487, 2020. [Online]. Available. https://www.sciencedirect.com/science/article/pii/S0920410519309088
2. O. J. Wilson, "Machine learning for well rate estimation: Integrated imputation and stacked ensemble modelling," Master's thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2020. [Online]. Available: https://dspace.mit.edu/handle/1721.1/132875
3. T. Al-Qutami, R. Ibrahim, and I. Ismail, "Hybrid neural network and regression tree ensemble pruned by simulated annealing for virtual flow metering application," in *Proceedings of the Conference on Signal and Image Processing Applications (ICSIPA)*, 2017. [Online]. Available: DOI: https://doi.org/10.1109/ICSIPA.2017.8120626

AUTHOR'S PROFILE

Mohammad Amir Ashraff is a Data Science Architect at BDB.ai with over 22 years of IT experience and more than 7 years specializing in data science, machine learning, and artificial intelligence. He has extensive hands-on expertise in designing, developing, and deploying ML, deep learning, NLP, and time-series solutions across industries, including oil & gas, automotive, energy, and fintech. At BDB.ai, he has worked as Product Owner for the Data Science Lab (DSLAB) and Automated Machine Learning platforms, leading large-scale industrial analytics initiatives. His professional experience includes developing predictive systems for oil production optimisation, automobile fuel-gauge accuracy, electrical fault prediction, and fraud detection. He holds a Bachelor's degree in Electrical Engineering from NIT Jamshedpur and a postgraduate qualification in AI and Machine Learning from Great Lakes Institute of Management, Hyderabad.