Comparative Evaluation of Prognostic Models for Medium-Term Pressure Prediction in Gas Transmission Units

Houda SARIH[1], Florent BRISSAUD[1], Khanh T. P. NGUYEN[2], Kamal MEDJAHER[2]

[1] *Natran R&I, 1-3 rue du Commandant d'Estienne d'Orves, 92390 Villeneuve-la-Garenne, France*
houda.sarih@natrangroupe.com
florent.brissaud@natrangroupe.com

[2] *Univ. Toulouse, UTTOP, LGP, 47 Av. d'Azereix, 65000 Tarbes, France*
thi-phuong-khanh.nguyen@uttop.fr
kamal.medjaher@uttop.fr

Abstract

Gas pressure reduction and delivery stations are critical components of gas transmission networks, where pressure drifts may signal early degradation or control malfunction. Despite the availability of measurements supporting data-driven Prognostics and Health Management (PHM), medium-term forecasting remains difficult due to non-stationarities, operator adjustments, and heterogeneous setpoints. This study investigates 5-day-ahead pressure forecasting for maintenance planning and early detection of evolutions preceding threshold exceedances. A robust preprocessing pipeline addresses irregular sampling, missing data, and level shifts under real operating conditions. Two interpretable models are evaluated: SARIMAX, for seasonal and exogenous effects, and LightGBM, for nonlinear dynamics with feature-importance analysis. Performance is assessed on a homogeneous subset of gas pressure measurement recorders using standard regression metrics. The results emphasize not only predictive capability but also interpretability, positioning transparent forecasting models as scalable and auditable building blocks for PHM in gas transmission networks.

Keywords: Prognostics and Health Management, time series forecasting; predictive maintenance; gas pressure monitoring; non-stationary time series

1. Introduction

Gas transmission and distribution networks are the backbone of modern energy infrastructure, and the regulatory integrity of gas pressure reduction and delivery stations is vital for their safe operation. The integrity of these assets is paramount because of safety, environment and economic issues. Consequently, the industry is increasingly shifting from reactive troubleshooting to PHM strategies that emphasize proactive failure avoidance through real-time health monitoring and predictive analytics (Chalgham, Wu, & Mosleh, 2020).

A core indicator of system health in these networks is gas pressure. Pressure drifts often constitute the earliest detectable indicators of degradation, control malfunctions, or emerging faults in components such as pipes, control valves, and compressors. However, accurately forecasting these drifts over the medium-term (5-day horizon) is complicated by strong non-stationarities, frequent operator-induced adjustments, and heterogeneous regulation setpoints (Brissaud, 2024). While high-frequency digital measurements have paved the way for data-driven PHM, the industry faces a growing tension between model complexity and operational transparency (Magdin, 2025).

Current literature demonstrates the strong performance of deep learning (DL) architectures for time series forecasting (Lim & Zohren, 2021), particularly in gas-related applications (Tian et al., 2024). For instance, the authors in (Wang et al., 2022) have proposed an end-to-end crack evolution forecasting framework for natural gas pipelines that combines advanced signal denoising with attention-based CNN–LSTM architectures, demonstrating the strong predictive capability for time-series failure anticipation. Similarly, a U-Net–BiLSTM–Attention architecture has been introduced for pressure regulation station load prediction in a recent study (Jie et al., 2025). In the same vein, a VMD–CNN–LSTM–Self-Attention framework has been developed for interval prediction of natural gas field station loads, integrating signal decomposition and confidence-based graded alarms to improve predictive reliability and safety-oriented decision support (Zhao et al., 2025). While effective, these DL models frequently lack the interpretability required for safety-critical maintenance decisions, where operators must understand the "why" behind a predicted failure (Lu & Frank Cheng, 2025). To address this, there is an increasing demand for models that balance high precision with algorithmic transparency.

In this light, classical statistical and tree-based ensemble models continue to demonstrate strong competitiveness in energy forecasting tasks, particularly when multiple seasonalities are present. For example, autoregressive and decomposition-based methods effectively model seasonal gas consumption patterns (Ding, Zhao, & Jin, 2023), while ensemble learning shows strong generalization in industrial time-series applications (Magdin, 2025). Accordingly, this paper evaluates two interpretable models for 5-day pressure forecasting: SARIMAX, which captures autoregressive, seasonal, and exogenous effects, and LightGBM, which models nonlinear interactions efficiently with feature-importance analysis. Both are integrated into a preprocessing pipeline handling irregular sampling, missing data, and non-stationary level shifts. Evaluation on a homogeneous subset of gas pressure measurement recorders (GPMR) establishes reliable baselines and shows that the framework supports auditable, scalable, and computationally efficient PHM for gas transmission networks.

The remainder of this paper is organized as follows. Section 2 presents the industrial context, dataset, and key challenges in pressure monitoring. Section 3 details the proposed framework, including preprocessing and the SARIMAX and LightGBM formulations. Section 4 discusses predictive performance, feature importance, and model behavior from an operational perspective. Finally, Section 5 summarizes the findings, PHM deployment implications, and future research directions.

2. INDUSTRIAL CONTEXT AND PROBLEM STATEMENT

2.1. Case study and data description

The present study uses measurements from **GPMRs** installed at gas pressure reduction and delivery stations (Fig. 1). These devices monitor regulated downstream pressure and transmit aggregated 5-minute values, together with hourly minimum and maximum indicators. The dataset covers July 2023 to September 2025 and includes over 235 million pressure observations from about 2,400 recorders. For methodological consistency, the analysis focuses on a homogeneous subset of 118 GPMRs with similar downstream regulation levels ($\approx$3.9 bar), identical station type, and at least 1.5 years of usable data. This subset was selected using a Gaussian Mixture Model to retain stations with comparable pressure distributions and dynamics.

Operator event logs record manual interventions, including regulator adjustments, maintenance actions, and operational procedures that may change the pressure regime. They are not used as explanatory variables; instead, they identify non-nominal periods excluded from training to prevent models from learning intervention-driven dynamics.

Figure 1. Gas pressure reduction and delivery station

Meteorological variables, especially air temperature, are used as optional exogenous regressors because they affect gas demand and indirectly influence pressure dynamics. For each station, observations from the nearest weather station are temporally aligned with pressure data to capture demand-related variations beyond pressure history.

2.2. Operational challenges

Although regulated, downstream pressure shows complex dynamics driven by seasonal demand, ambient temperature, operator interventions, setpoint changes, telemetry gaps, and possible equipment aging. A major industrial challenge is the lack of reliable labels for degradation onset: pressure variations may reflect normal operations rather than faults, while maintenance can generate extreme but non-failure-related values. Since intervention logs are often incomplete or misaligned with measurements, separating operational changes from true anomalies remains difficult, making supervised anomaly detection poorly suited.

To address these limitations, this work adopts a prediction-based monitoring strategy consistent with condition-based maintenance. The task is formulated as a 5-day multistep forecasting: *(i) learning nominal pressure dynamics from historical data, (ii) predicting future trajectories, and (iii) interpreting sustained forecast–observation deviations as signs of emerging drift*. This avoids explicit anomaly labels and reframes degradation detection as deviation from expected behavior, which is well-suited to industrial settings with rare failures and frequent operational perturbations. Within this scope, the contribution of this paper is *not to design a full anomaly decision system*, but to **investigate interpretable forecasting baselines** suitable for pressure monitoring under real operational constraints.

3. METHODOLOGICAL FRAMEWORK

Figure 2 summarizes the methodological framework used for medium-term downstream gas pressure forecasting.

First, heterogeneous data are collected from GPMRs, opera-

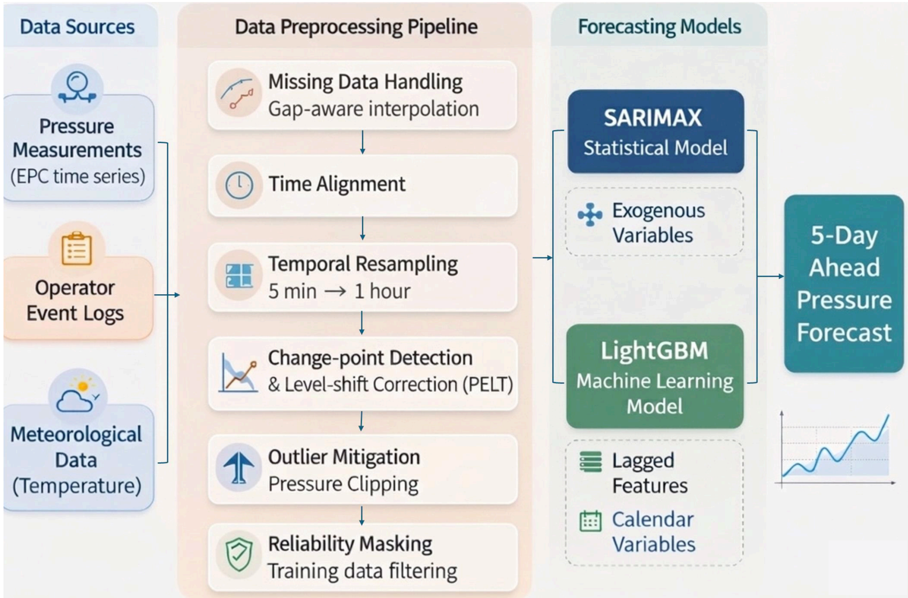


Figure 2. Workflow of the proposed forecasting framework

tor intervention logs, and meteorological variables, providing complementary information on network behavior and operating context. A unified preprocessing pipeline then converts raw measurements into learning-ready time series by applying gap-aware interpolation, temporal alignment, hourly resampling, level-shift correction via change-point detection, outlier mitigation, and reliability masks to remove unstable periods. Finally, the processed data are used to train two complementary forecasting models: a statistical SARIMAX model capturing seasonal and exogenous effects, and a LightGBM model that leverages nonlinear relationships through lagged pressure variables and calendar features. Both models produce five-day multistep forecasts, matching the operational monitoring horizon. The framework components are detailed in the following subsections.

3.1. Handling missing data

Missing observations are recurrent in industrial telemetry and mainly arise from two mechanisms: **prolonged interruptions** due to station shutdowns or sensor unavailability, and **short gaps** caused by delayed or failed telemetry transmissions. To handle them while preserving physical plausibility, a **gap-aware interpolation strategy** is applied according to the length of consecutive missing segments. Each gap is identified by run-length encoding of the null indicator; gaps shorter than or equal to $G_{\max}$ are interpolated, whereas longer gaps are kept missing.

For the original five-minute pressure measurements, $G_{\max} = 12$, corresponding to a maximum interpolated duration of one hour. Short gaps are filled by linear spline interpolation using valid neighboring observations, whereas longer gaps are kept missing to avoid introducing artificial dynamics over extended interruptions. Since the selected 118 GPMRs contain at least 1.5 years of usable data over an approximately 2.1-year observation period, interpolation affects only a limited fraction of the dataset, and training remains based predominantly on observed values.

3.2. Time alignment and resampling

All data sources are synchronized on a common timestamp axis to ensure temporal consistency. Pressure measurements, meteorological observations, and operator logs are converted to a common datetime format and aligned with the reference pressure timeline. Meteorological variables are merged with each GPMR series using the linked weather stations, while operator events are projected onto the same temporal grid to identify the corresponding pressure observations.

Pressure measurements, originally recorded every five minutes, are then resampled to an hourly resolution by window averaging. This choice ensures consistency with hourly meteorological data, remains adequate for capturing gradual pressure drifts, and reduces both the forecasting horizon length and computational burden while filtering high-frequency consumption noise.

However, temporal aggregation inevitably entails some information loss and potential aliasing effects. To quantify this impact, the **Fraction of Variance Unexplained (FVU)** is computed by comparing the original pressure signal with the hourly aggregated signal reconstructed at five-minute resolution:

$$\mathrm{FVU} = \frac{\mathrm{MSE}(f)}{\mathrm{Var}(Y)}$$

where $\mathrm{MSE}(f)$ is the mean squared reconstruction error and $\mathrm{Var}(Y)$ the local variance of the original signal computed over a one-week sliding window. Across the studied GPMR sample, the average FVU is approximately 16%, indicating that the hourly representation preserves most signal variability while filtering high-frequency fluctuations.

3.3. Change-point detection and level-shift correction

Pressure time series may exhibit abrupt level shifts caused by operator setpoint adjustments or station configuration changes. These regime changes introduce structural breaks that can degrade forecasting performance by violating the assumption of locally stable statistical properties. To mitigate this effect, change points are detected using the **Pruned Exact Linear Time (PELT) algorithm** implemented in the Python *Ruptures* library. A linear kernel is used with a *minimum segment length of 360 samples and a penalty parameter of 0.5*, allowing an unknown number of structural breaks to be identified.

For each detected rupture, the shift magnitude is estimated by comparing local averages before and after the change point using two empirically selected windows, $[t-40, t-5]$ and $[t+5, t+40]$. These empirically selected windows provide stable local means while excluding the immediate rupture neighborhood to avoid transient effects. The difference between the corresponding two mean values defines the estimated jump. Detected shifts are then processed in reverse chronological order and cumulatively subtracted from subsequent observations, yielding a corrected series in which regime-induced offsets are removed while short-term dynamics are preserved.

The effect of this correction is illustrated in Fig. 3, which shows examples of pressure series before (blue line) and after the level-shift adjustment (orange line). The procedure effectively removes abrupt regime-induced offsets while preserving the local variability and short-term dynamics of the signal. As a result, the corrected series exhibit more consistent statistical properties across time. This transformation allows the models to focus on gradual pressure drifts associated with regulator behavior rather than on discrete operator interventions.

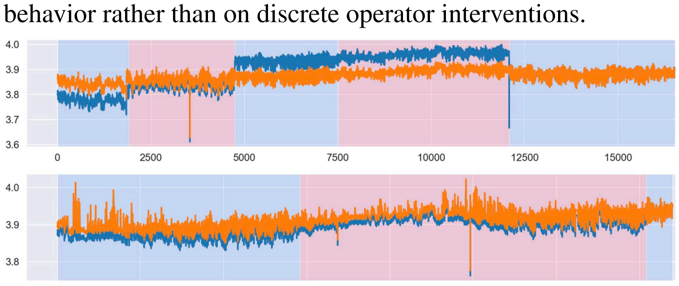


Figure 3. Illustration of level-shift removal in pressure (in bar)

3.4. Outlier mitigation and reliability masking

Pressure time series occasionally contain extreme values or short spikes associated with *station purging, maintenance operations, or transient telemetry artifacts*. Because these events are rare and poorly documented, explicitly modeling them would introduce noise and degrade forecasting performance. To mitigate their impact, *outlier observations are handled using a clipping strategy*. The admissible pressure range is defined as ± 6 *times a locally estimated standard deviation over a sliding one-month window*. This deviation is estimated using the **Optimal Quantile Absolute Deviation (OQAD)** method, which provides a robust scale estimate in the presence of outliers. Values exceeding this range are clipped to the nearest boundary rather than removed, preserving temporal continuity while limiting the influence of extreme observations. In practice, only a very small fraction of observations are affected by this procedure, confirming the rarity of such events. As illustrated in Fig. 4, clipping effectively limits extreme spikes while preserving the overall structure and normal operating dynamics of the pressure series, thereby improving the robustness of the forecasting models.

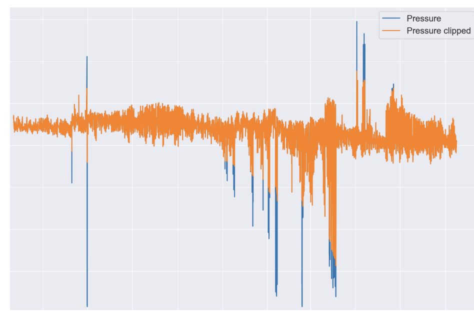


Figure 4. Illustration of outlier removal in pressure (in bar)

In addition to outlier mitigation, a **binary reliability mask** m_t is constructed to *exclude unreliable timestamps during model training*. The mask combines data-quality and operational criteria. First, temporal dilation is applied around filtered observations, such as clipped values or interpolated segments, using *a centered rolling maximum with a window* of ± 3 hours. Second, *intervention logs are used to mask periods surrounding operator actions*, with an exclusion window of approximately five hours around each event. This conservative strategy limits contamination of the nominal training set by poorly documented transients, since anomaly labels are incomplete and the true pressure state cannot be reliably reconstructed. More data-efficient alternatives, such as Kalman filtering or robust smoothing, could preserve additional samples but require stronger assumptions on the underlying dynamics and are left for future work. When reliable event annotations become available, such transients could be retained as a dedicated test subset for residual-based anomaly detection. Finally, to ensure an unbiased evaluation, the mask is forced to zero over the final forecasting horizon, keeping this period fully available for testing.

3.5. Forecasting models

SARIMAX model. SARIMAX is implemented at the *individual GPMR level*, with one model fitted for each series and forecasting cycle. Because SARIMAX relies on the recent autoregressive structure of the signal, training is performed on a sliding window rather than the full historical series. Temperature is used as the only exogenous variable, and the seasonal period is fixed to $s = 24$ to capture daily seasonality in hourly data.

Fig 5 presents the SARIMAX training and evaluation workflow. Its hyperparameters are optimized using **Optuna**, which employs a *Tree-structured Parzen Estimator (TPE)* sampler, a Bayesian optimization method that sequentially explores the parameter space based on previous trials. The optimized parameters include the non-seasonal order (p, d, q), the seasonal order $(P, D, Q, 24)$, the trend specification (`n` or `c`), and the training-window length. The search space is defined as $p \in \{1, 2, 3\}$, $d \in \{0, 1\}$, $q \in \{1, 2, 3\}$, $P, D, Q \in \{0, 1\}$, with training windows ranging from 30 to 180 days. Each configuration is evaluated using a fixed temporal split, where the *last 5 days are reserved for testing, the preceding 5 days for validation, and the remaining portion of the window for training*. The objective minimized by Optuna is the average validation RMSE across GPMRs.

After optimization, the best configurations are re-evaluated on all GPMR. For each configuration, the model for each GPMR is first trained on the training subset to forecast the validation horizon, then re-fitted on the combined train and validation data to generate the final 5-day test forecast. This procedure provides an interpretable statistical benchmark while accounting for seasonality and meteorological effects, albeit with a higher computational cost due to repeated model re-estimation.

LightGBM model. LightGBM is implemented as a ***global forecasting model*** trained on the concatenation of all GPMR samples, where the target corresponds to the **next-hour pressure**. As illustrated in Fig. 6, the workflow first constructs a feature representation combining *exogenous variables* and *autoregressive descriptors* of the pressure signal. The exogenous block includes temperature and calendar variables (`hour`, `week`, `weekday`, `month`, `day`, `days_in_month`), together with the GPMR identifier. The autoregressive block consists of **16 lagged pressures** $[1, 2, 3, 4, 5, 6, 7, 12, 24, 36, 48, 72, 96, 120, 144, 168]$ and **rolling statistic features (RSF)** computed over *24-hour* and *168-hour* windows. These RSF summarize recent pressure dynamics through the *mean, median, standard deviation, minimum, and maximum* of the preceding observations.

For each GPMR, a feature table is constructed by merging pressure and exogenous variables, then generating lag and rolling features. The current pressure value is excluded from predictors to prevent leakage. The resulting dataset contains **46 features**, including meteorological and calendar descriptors, the GPMR identifier, 16 lagged variables, and 10 rolling statistics. The GPMR identifier is used only as a categorical grouping variable in the pooled LightGBM model to capture station-specific heterogeneity. It is not considered a physical feature and is excluded from the physical interpretation of feature importance to avoid bias. Data are split chronologically into **training, validation, and test sets**, with the last *5 days* reserved for testing and the preceding *5 days* for validation. Because the maximum lag equals *168 hours*, the first *168 training samples* are discarded to preserve feature-label alignment.

LightGBM model is trained once on the pooled training set using a gradient-boosted decision tree objective (`l2`). During validation and testing, forecasts are produced using a **recursive multi-step strategy**: *at each step, lag and rolling features are reconstructed from the available history, combining observed past values with previously predicted ones, while exogenous variables are directly injected.* This approach enables **5-day ahead forecasting** using a single one-step model while remaining consistent with deployment conditions.

4. RESULTS AND OPERATIONAL DISCUSSION

The results reported in this section were obtained using Python 3.13.3 on macOS Sequoia 15.7.3, with an Apple M4 Pro processor, 12 CPU cores, and 24 GB of unified memory. Since no deep learning reference model is included, the reported errors should be interpreted as baseline performance levels for interpretable models, rather than as evidence of general competitiveness against more complex architectures.

SARIMAX model. For SARIMAX, the hyperparameter search was performed globally over the homogeneous GPMR subset: each candidate configuration was evaluated on all available GPMRs, and the objective minimized the mean validation RMSE across stations. The configurations reported in Tab. 1 therefore correspond to the best aggregate-performing settings, rather than the most frequently selected configurations from independent per-GPMR optimizations. The first configuration corresponds to $(p, d, q) = (3, 1, 1)$ with seasonal order $(1, 0, 1)$ and a training window of 3600 samples ($\approx 150\,days$). The relatively autoregressive order ($p = 3$) indicates that pressure dynamics depend on 3 recent observations, reflecting the *short-term inertia of gas networks and the stabilizing effect of regulator control loops*. The presence of a moving-average term ($q = 1$) captures *transient disturbances and stochastic fluctuations related to consumption variability*. The first-order differencing ($d = 1$) removes slow drifts in the signal, which may arise from *gradual demand variations or operator adjustments of pressure setpoints*. Finally, the seasonal autoregressive and moving-average components $(P, Q) = (1, 1)$ model *recurring daily operational patterns* in pressure dynamics.

The two alternative configurations, $(2, 1, 2)$ and $(1, 0, 3)$, show a similar balance between autoregressive and moving-average components, confirming that *pressure evolution results from both persistent system dynamics and short-term stochastic perturbations*. The selected training windows (ranging from 90 to 180 days) suggest that *reliable parameter*

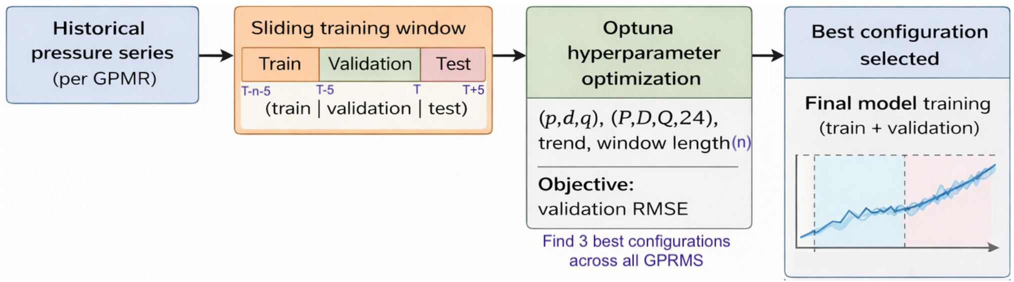


Figure 5. SARIMAX training and evaluation workflow

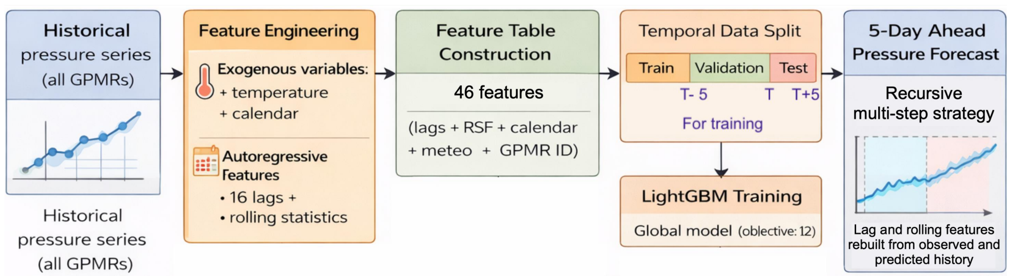


Figure 6. LightGBM training and evaluation workflow

Table 1. Best SARIMAX configurations identified by Optuna across all GPMRs.

Rank	(p, d, q)	(P, D, Q, s)	Trend	Train window
1	(3,1,1)	(1,0,1,24)	c	3600
2	(2,1,2)	(1,1,1,24)	c	4320
3	(1,0,3)	(0,1,0,24)	c	2160

estimation requires capturing multiple operational cycles. Finally, the inclusion of a constant trend term (`trend = c`) allows the model to account for a non-zero baseline level in the pressure signal, *reflecting the regulated operating setpoint around which short-term fluctuations occur*.

Following the proposed workflow in Fig. 5, each GPMR time

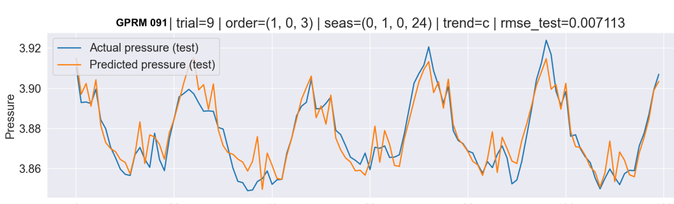


Figure 7. Pressure forecast (in bar) using SARIMAX

series is trained using the 3-best SARIMAX configurations, and the model achieving the lowest validation error is retained to produce the final forecast. Across all GPMRs, the resulting **mean test RMSE** ≈ 11.5 mbar, indicating that the model captures the main pressure dynamics with relatively acceptable deviations compared to the typical operating pressure range.

Figure 7 illustrates the 5-day forecast for a representative GPMR (ID 091), achieving $RMSE \approx 7.1$ mbar. The model accurately reproduces the amplitude and timing of daily pressure oscillations. Small deviations near sharp extrema reflect the typical accumulation of errors in recursive multi-step forecasting.

LightGBM model. Unlike SARIMAX, which fits an individual model for each GPMR, LightGBM is trained as a *single global model* using the concatenation of the 118 GPMR time series. This strategy enables the model to learn shared pressure dynamics across stations while exploiting a richer feature set composed of *lagged pressures, rolling statistics, meteorological, and calendar features*. Across all GPMRs, the resulting **mean test RMSE** is approximately 13.1 mbar. Although slightly higher than the SARIMAX benchmark (≈ 11.5 mbar), this error remains moderate relative to the typical operating pressure range and confirms the capability of the global model to reproduce the main pressure dynamics across heterogeneous stations.

Fig. 8 presents an illustrative forecast for the same repre-

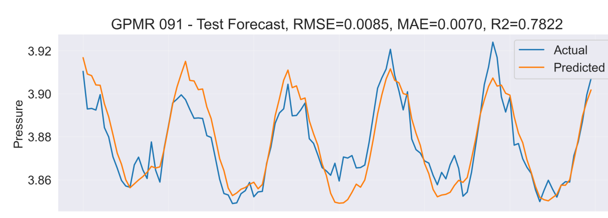


Figure 8. Pressure forecast (in bar) using LightGBM

sentative GPMR (ID 091) over the 5-day prediction horizon.

The prediction achieves $RMSE \approx 8.5$ mbar, indicating that the model captures most of the temporal variability of the pressure signal. The predicted trajectory follows the observed daily oscillations and overall trend, although the forecasts appear smoother than those obtained with SARIMAX (Fig. 7 vs. Fig. 8). This smoothing effect is typical of global machine-learning models and explains the slightly higher RMSE, as sharp local variations tend to be attenuated.

To better interpret the behavior of the LightGBM model,

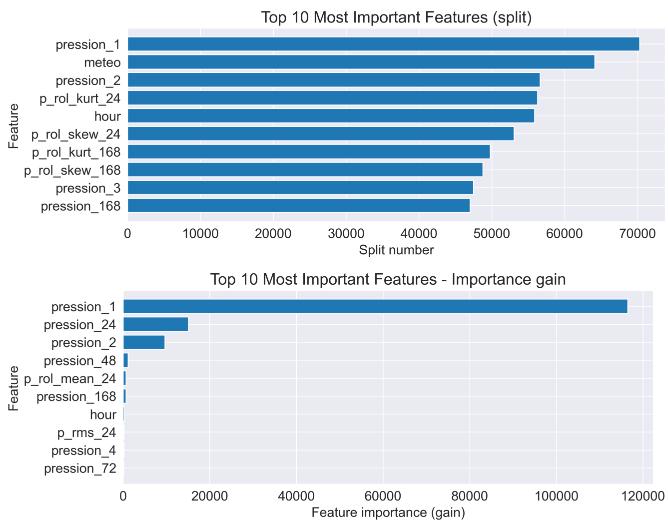


Figure 9. LightGBM feature importance analysis

Fig. 9 reports the ranking of the most influential predictors according to two complementary importance metrics: *split number* and *gain importance*. The *split number* measures how frequently a feature is used to partition the decision trees, indicating variables that repeatedly contribute to the model structure. In contrast, *gain importance* quantifies the reduction of the loss function obtained when a feature is used for a split, thus highlighting the predictors that contribute most to improving forecasting accuracy.

According to the split-based ranking, several predictors contribute regularly to the model decisions, including short-term pressure lags (`pression_1`, `pression_2`, `pression_3`), meteorological conditions (`meteo`), rolling statistical descriptors of the pressure signal (e.g., `p_rol_kurt_24`, `p_rol_skew_24`), and the calendar variable `hour`. This distribution indicates that the model systematically combines recent system states, demand-related meteorological drivers, and daily operational cycles to describe pressure dynamics. In particular, the strong presence of rolling statistics suggests that the model captures short-term regime changes in pressure variability, which may correspond to transient demand peaks or local network adjustments.

The gain-based ranking provides a complementary perspective. It shows that the immediate pressure history, especially `pression_1`, dominates the predictive power of the model, with a significantly larger contribution to error reduction than other variables. Additional lags such as `pression_24` and `pression_2` also play an important role, indicating that the model relies heavily on the persistence of pressure states over short and daily horizons. In comparison, meteorological and calendar features appear less influential in terms of gain, suggesting that they mainly act as contextual variables that refine predictions rather than driving them directly.

These observations are consistent with the SARIMAX configurations reported in Tab. 1. The best-performing models systematically include autoregressive components with several recent lags (e.g., $p = 1, 2, 3$) as well as a seasonal component with period $s = 24$. This structure reflects the same two dominant mechanisms highlighted by the LightGBM importance analysis: short-term persistence of pressure dynamics and a strong daily periodicity. Additionally, the LightGBM approach therefore offers a different trade-off between accuracy and scalability. While the statistical model slightly outperforms the global model in terms of average RMSE, LightGBM provides a single global model, reducing the overall training time. From a deployment viewpoint, LightGBM is particularly attractive because, once trained, it provides fast CPU-based inference with limited memory requirements, making it suitable for large-scale monitoring of multiple GPMRs. SARIMAX offers strong station-level interpretability, but its per-series fitting strategy is less scalable when the number of monitored recorders increases.

From an **operational perspective**, the SARIMAX and LightGBM results confirm that downstream pressure evolution is mainly governed by the recent pressure trajectory itself, reflecting both the inertia of the gas transmission system and the feedback regulation mechanisms of pressure control devices. Exogenous and calendar variables nevertheless provide complementary information for anticipating recurrent demand cycles and seasonal consumption patterns. For a future fault detection stage, the decision boundary could be calibrated from the distribution of nominal forecast residuals using validation-based quantiles or prediction intervals, with alarms triggered only when deviations persist over consecutive time steps. These findings open several directions for further development. Sequence-based architectures, such as LSTM could better capture short-term pressure dynamics, while attention-based or transformer models may learn longer-range dependencies and multi-scale seasonal patterns, such as daily or weekly demand cycles. Future work could also integrate heterogeneous inputs and physical knowledge by combining pressure history, meteorological, and calendar variables within hybrid or physics-informed architectures, and by exploiting spatio-temporal models, such as graph neural networks, to capture pressure propagation across interconnected stations.

5. CONCLUSIONS AND PERSPECTIVES

This study investigated the problem of short-term downstream pressure forecasting in gas transmission networks using two complementary modeling approaches: the statistical SARIMAX model and the tree-based ensemble method LightGBM. Both models are integrated within a dedicated preprocessing workflow designed to address practical challenges of industrial monitoring data, including irregular sampling, missing values, and non-stationary level shifts. The results show reliable predictions over a 5-day horizon, with SARIMAX achieving slightly lower average RMSE, while LightGBM demonstrates strong robustness and scalability through a single global model trained across all stations.

From an operational standpoint, the study highlights the importance of combining interpretable statistical and machine-learning models for industrial monitoring tasks. Such approaches remain competitive in energy forecasting applications and provide transparent mechanisms to analyze the drivers of pressure dynamics while maintaining computational efficiency for large monitoring infrastructures.

Future work could extend the proposed framework through systematic comparisons with advanced data-driven models and reference baselines, including LSTM, DLinear, persistence, and seasonal naïve forecasting, in order to better quantify the relative predictive gain of interpretable approaches. Integrating physics-informed learning strategies and spatio-temporal representations may also help capture multi-scale temporal patterns and spatial dependencies across interconnected stations, thereby improving forecasting robustness for large-scale gas transmission networks.

ACKNOWLEDGMENT

The authors would like to acknowledge that some preliminary results were developed in the context of a Master's internship carried out at Natran R&I by Mr. Maxyme Soudani. The authors thank him for his contribution to the initial exploratory analyses that supported this study.

REFERENCES

Brissaud, F. (2024, May). Predictive maintenance to prevent drifts applied to gas pressure reduction and delivery stations. In *International gas research conference (igrc 2024).* Banff, Canada. (Presented May 13–16, 2024)

Chalgham, W., Wu, K.-Y., & Mosleh, A. (2020, December). System-level prognosis and health monitoring modeling framework and software implementation for gas pipeline system integrity management. *Journal of Natural Gas Science and Engineering*, *84*, 103671.

Ding, J., Zhao, Y., & Jin, J. (2023, May). Forecasting natural gas consumption with multiple seasonal patterns. *Applied Energy*, *337*, 120911.

Jie, H., Yuan, C., Liu, H., Liu, Q., Qiu, X., Zhang, M., & Zhao, M. (2025, April). Deep Learning-Based Signal Processing and Prediction for Gas Pressure Regulating Stations. *Journal of Physics: Conference Series*, 012055.

Lim, B., & Zohren, S. (2021, April). Time Series Forecasting With Deep Learning: A Survey. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *379*(2194).

Lu, H., & Frank Cheng, Y. (2025, September). Artificial Intelligence in Energy Pipelines: Opportunities and Risks. *Engineering*.

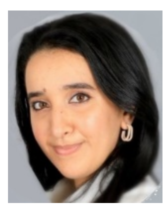
Magdin, M. (2025, December). Literature review: Current trends and advances in the use of artificial intelligence for ensuring the safety and efficiency of gas pipeline operations. *Results in Engineering*, *28*, 107309.

Tian, N., Shao, B., Bian, G., Zeng, H., Li, X., & Zhao, W. (2024, March). Application of forecasting strategies and techniques to natural gas consumption: A comprehensive review and comparative study. *Engineering Applications of Artificial Intelligence*, *129*, 107644.

Wang, B., Guo, Y., Wang, D., Zhang, Y., He, R., & Chen, J. (2022, December). Prediction model of natural gas pipeline crack evolution based on optimized DCNN-LSTM. *Mechanical Systems and Signal Processing*, *181*, 109557.

Zhao, W., Shao, B., Tian, N., Zhang, W., Zhao, X., & Wu, S. (2025, February). Early warning study of field station process safety based on VMD-CNN-LSTM-self-attention for natural gas load prediction. *Scientific Reports*, *15*(1), 6360.

BIOGRAPHIES



Houda SARIH is an RD Activity Leader and Research Engineer at NaTran RI (GRTgaz), France, where she has been contributing to research and innovation activities in gas network reliability and maintenance engineering since 2024. She received her Ph.D. in Industrial Engineering from the National Polytechnic Institute of Toulouse (INP Toulouse) in 2020. She previously worked as a reliability and safety engineer in the railway sector, focusing on functional safety analyses and failure risk assessment. Her research focuses on functional safety and data-driven predictive maintenance.

Florent BRISSAUD is a Lead Research Engineer Expert at NaTran, France, where he has worked since 2018 as a RD Manager in reliability, functional safety, and industrial asset management. He received his Ph.D. in Systems Modelling and Dependability from the University of Technology of Troyes in 2010, as well as an Executive Master from Université Paris Dauphine-PSL focusing on artificial intelligence and data science. Prior to joining NaTran, he worked for nearly a decade as a consultant in the energy and industrial sectors. His research interests include reliability modelling, maintenance optimization, and the integration of artificial intelligence into industrial asset management.

Khanh T. P. Nguyen is a Full Professor at the University of Technology Tarbes Occitanie Pyrénées, France, since 2025 and a member of the Production Engineering Laboratory (LGP) since 2017. She received her Ph.D. in Automation and Production Engineering from École Centrale de Nantes in 2012. She was a Postdoctoral Fellow at IFSTTAR from 2013 to 2015, an Assistant Professor at the University of Technology of Troyes from 2016 to 2017, and an Associate Professor at the National Engineering School of Tarbes (ENIT) from 2017 to 2025. Her research focuses on artificial intelligence for predictive maintenance and Prognostics and Health Management.

Kamal Medjaher is a Full Professor at the University of Technology Tarbes Occitanie Pyrénées (UTTOP), France, and member of the Production Engineering Laboratory (LGP). He received the Ph.D. degree in control and industrial computing from the University of Lille 1, in 2005. He was an Associate Professor at the National Institute of Mechanics and Microtechnologies, Besançon, and FEMTO-ST Institute, from 2006 to 2016. His research works include prognostics and health management of industrial systems and predictive maintenance.