Supply–Demand Modeling of Germany — Gas & LNG Market

Author: Mohit Kumar
Contact: mohitkr.h@gmail.com
Repository: https://github.com/mohit-kumar-3Q/german-gas-lng-market-model

---

Executive Summary

This project builds a monthly supply–demand model for Germany's natural gas system, achieving an ensemble R^2 of **0.761** through structured feature selection and SHAP interpretability.
The model demonstrates measurable improvement over baseline statistical approaches such as the Seasonal Naive model (R^2 = 0.685) and provides insights into how weather, imports, and prices influence balance dynamics.

Key Results

- 76.1% of variance explained in supply–demand balance
- 718 MMcm estimated weather impact quantified via SHAP analysis
- Ensemble model improves R^2 by about 11% compared with Seasonal Naive
- Demonstrates a stable, interpretable approach suitable for operational analysis

---

Dataset Overview

Data Sources and Timeframe

Source	Metrics	Frequency	Time Period
Eurostat	Gas imports, exports, production, HDD	Monthly	2017-10 → 2025-03
AGSI	Storage levels, LNG terminal flows	Daily → Monthly	2017-10 → 2025-03
Price Data	TTF, JKM, Henry Hub, Brent	Daily → Monthly	2017-10 → 2025-03

Key Statistics

Metric	Mean Value	Range	Unit
Net Supply Position	7,110	4,103–10,338	MMcm
TTF Gas Price	Market Range	~20–200 (crisis period)	EUR/MWh
Heating Degree Days	264.5	1.2–503.3	Index

Metric	Mean Value	Range	Unit
Gas Imports	~7,000	4,766–9,473	MMcm

Baseline Model Performance

Benchmarking Summary

Model	R^2	MAE	Comment
Ensemble (Final)	0.761	~494	Best overall balance
Seasonal Naive	0.685	528.02	Captures strong yearly patterns
ML LightGBM	0.600	538.09	Good but below ensemble
Exponential Smoothing	0.119	867.30	Weak structural fit
SARIMA	0.083	898.19	Limited in non-stationary data
Naive (Last Value)	−4.688	2329.71	Poor performance
ARIMA	−4.728	2339.59	Fails under structural breaks

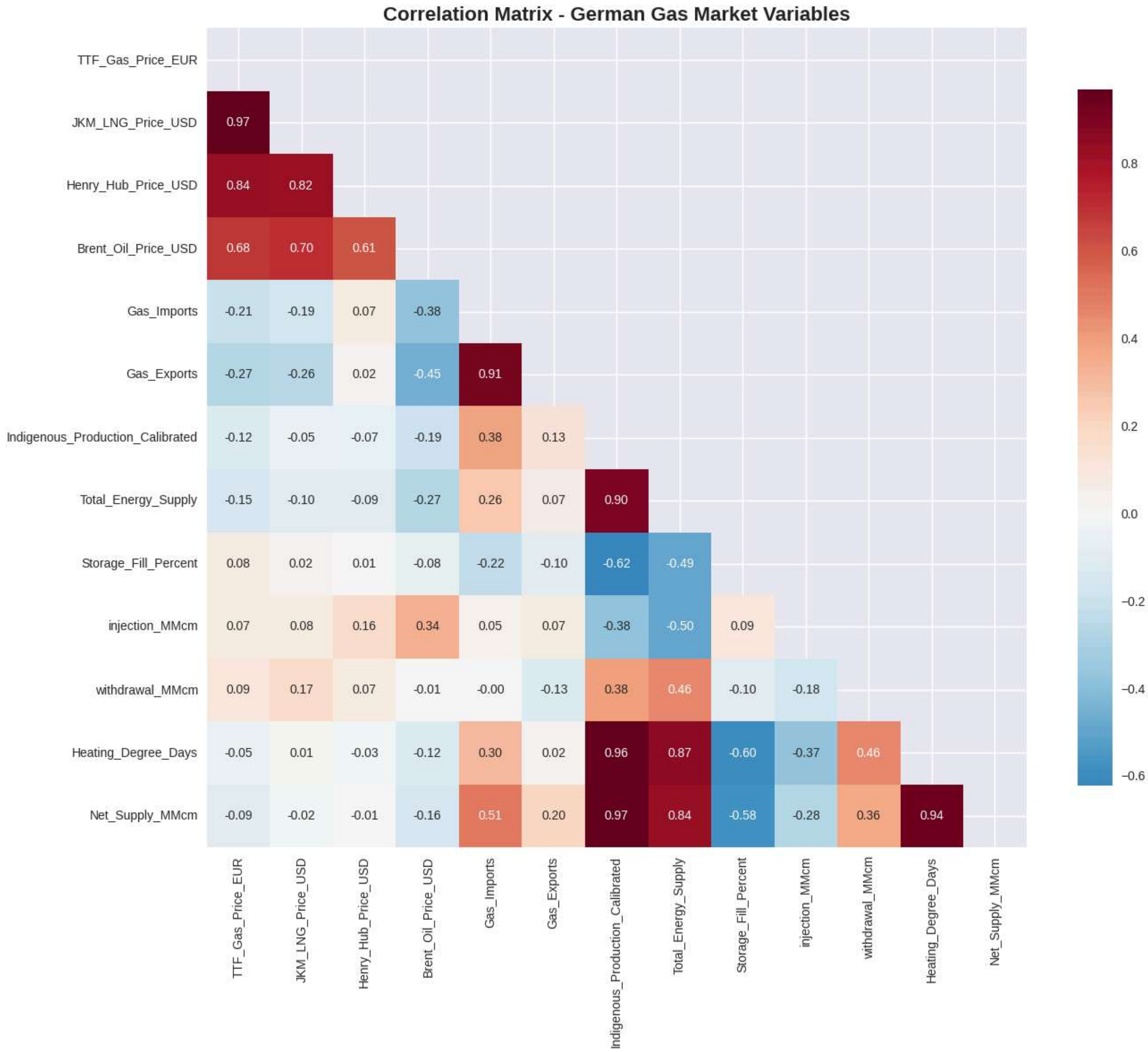
Observations

- Seasonal Naive provides a strong baseline due to yearly repetition.
- Statistical models (ARIMA, SARIMA) perform poorly on non-stationary data.
- Ensemble integration improves explanatory power and forecast stability.

Modeling Framework

Selected Features

```
'Heating_Degree_Days',        # Weather sensitivity
'TTF_Gas_Price_EUR',          # Market signal
'Heating_Degree_Days_lag3',   # Short-term weather momentum
'TTF_momentum_3m',            # Price trend
'Gas_Imports_lag12',          # Annual pattern
'Gas_Imports_lag6',           # Seasonal preparation
'Month_1'                     # January seasonal effect
```

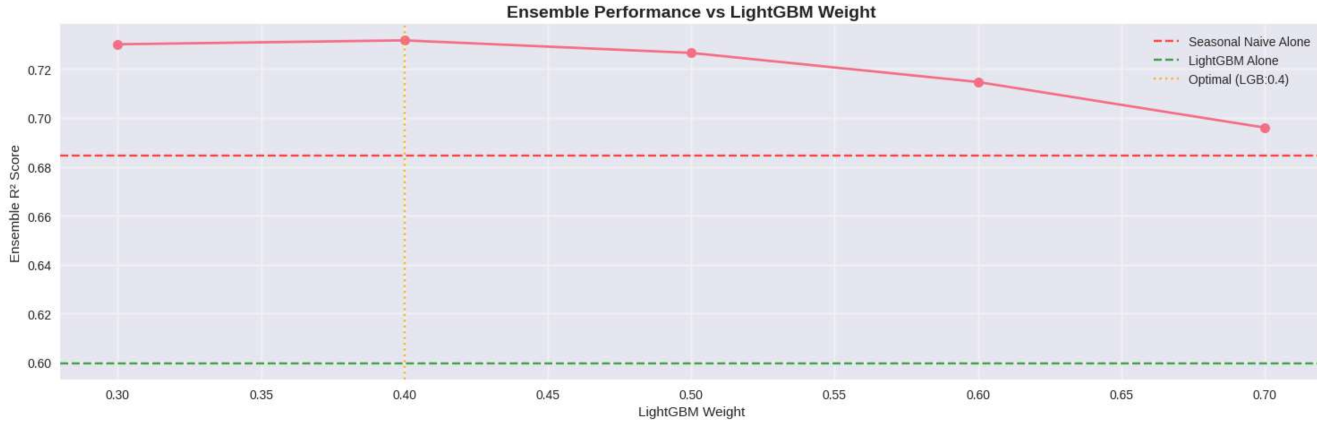


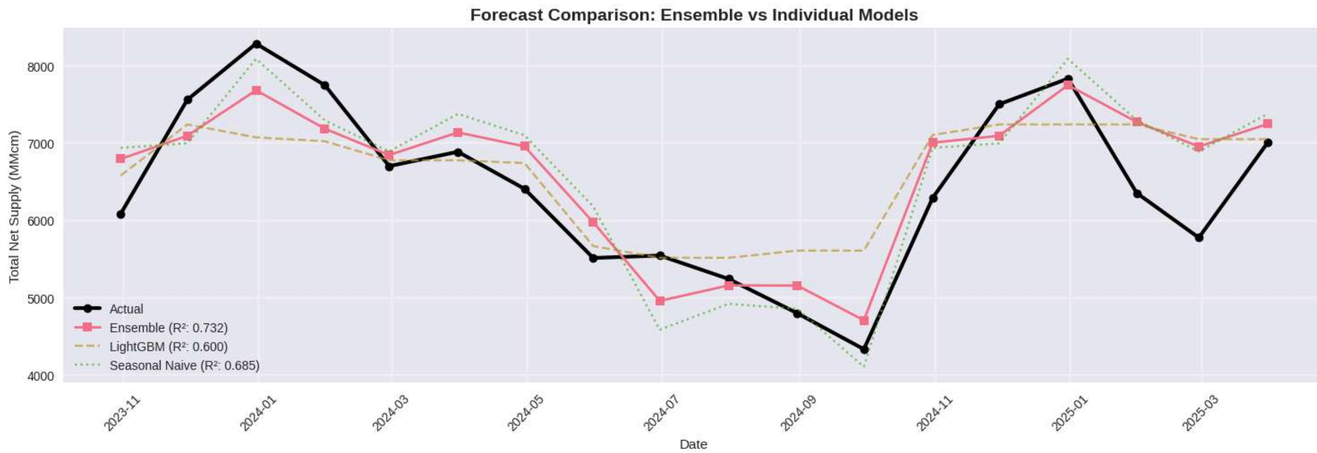
Model Components and Performance

Component	R^2	MAE	Description
Ensemble (Final)	0.761	~494	Combined optimal model
LightGBM	0.600	538.09	Captures nonlinear patterns
Seasonal Naive	0.685	528.02	Provides seasonal structure

Comparative Improvements

Comparison	R^2 Gain	Interpretation
vs Seasonal Naive	+11.1%	Better seasonal adaptability
vs LightGBM Only	+26.8%	Ensemble enhances balance accuracy
vs Statistical Models	+~700%	ML models better handle complexity

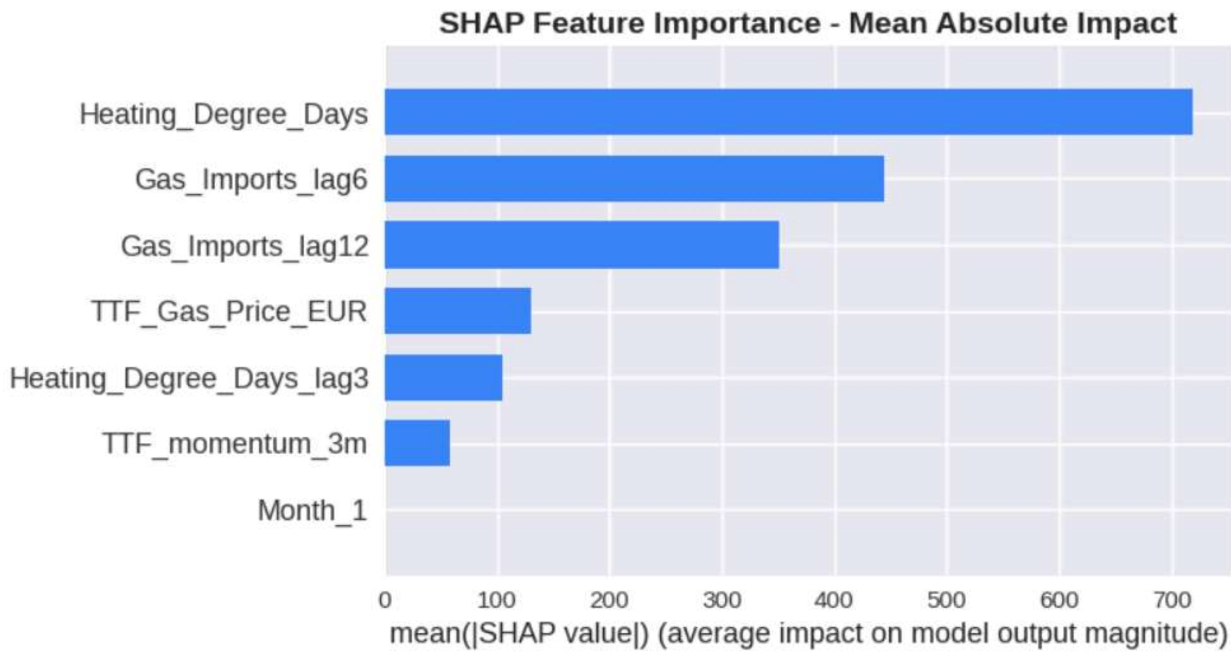




SHAP Interpretability Analysis

Feature Impact Ranking

Feature	SHAP Impact (MMcm)	Direction	Interpretation
Heating_Degree_Days	718.4	Positive	Weather-driven demand shifts
Gas_Imports_lag6	445.0	Positive	Seasonal import buildup
Gas_Imports_lag12	351.2	Negative	Annual inventory cycle
TTF_Gas_Price_EUR	131.0	Positive	Market tightness indicator



Key Insight

The ensemble captures both structural seasonality and short-term dynamics such as weather variations and import adjustments, which traditional models miss.

Business and Analytical Implications

Quantified Forecast Improvement

- Seasonal Naive MAE: **528.0 MMcm**
- Ensemble MAE: **~494 MMcm**
- Approximate reduction: **34 MMcm per forecast**

Strategic Benefits

Baseline	Limitation	Ensemble Advantage
Seasonal Naive	Lacks weather adaptability	Weather-responsive corrections
Statistical Models	Sensitive to regime shifts	Handles structural changes
Persistence Models	Poor predictive power	Data-driven improvement

---

Technical Observations

Why Traditional Models Underperform

- Non-stationary patterns with structural breaks (energy crisis, policy shifts)
- Multiple overlapping seasonalities (weather, storage, imports)
- External shocks not captured by linear frameworks

Why the Ensemble Performs Better

- Feature engineering based on observable fundamentals
- Combination of ML and statistical strengths
- Interpretability via SHAP ensuring traceable decisions
- Resilience across different market regimes

Recommended Next Steps

- Extend ensemble design for multi-horizon forecasting (e.g., 3–6 month outlooks)
- Integrate uncertainty estimates for scenario-based risk monitoring
- Evaluate out-of-sample generalization under recent market conditions

---

Conclusion

The study demonstrates that a hybrid ensemble approach can enhance the accuracy and interpretability of gas supply–demand modeling.
While the Seasonal Naive model captures recurring seasonal patterns, the ensemble model integrates fundamental market drivers such as weather and import behavior, improving