Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

VP-OilEx-STFN: Structural-Gated Mixed-Frequency Nowcasting for Oil Export Prediction in Norway and Thailand

Uparittha Intarasat
Faculty of Science and Technology, Rajamangala University of Technology Suvarnabhumi
Nassamon Bootwisas
nassamon.b@rmutsb.ac.th
Faculty of Science and Technology, Rajamangala University of Technology Suvarnabhumi
Pasin Marupanthorn
Faculty of Science, Maejo University

Research Article

Keywords: oil export nowcasting, mixed-frequency learning, structural deep forecasting, energy trade analytics

Posted Date: July 7th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9995316/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

VP-OilEx-STFN: Structural-Gated Mixed-Frequency Nowcasting for Oil Export Prediction in Norway and Thailand

Uparittha Intarasat[1], Nassamon Bootwisas[1*], Pasin Marupanthorn[2]

[1]Division of Computational Science, Faculty of Science and Technology, Rajamangala University of Technology Suvarnabhumi, Phra Nakhon Si Ayutthaya, 13000, Thailand.

[2]Department of Mathematics, Faculty of Science, Maejo University, 63 Moo 4, Nong Han, San Sai, Chiang Mai, 50290, Thailand.

*Corresponding author(s). E-mail(s): nassamon.b@rmutsb.ac.th;

Abstract

Official oil-export statistics are released with delay, yet governments, central banks, firms, and researchers need timely signals about export revenue, external balances, and energy-market stress. This paper develops a real-time mixed-frequency nowcasting framework for monthly oil-export prediction and studies two structurally different exporters: Norway, where exports are mainly governed by upstream crude and liquids production, and Thailand, where petroleum exports are shaped by refinery margins, crude inputs, domestic demand, and regional product markets. We propose the Vintage-Patch Oil Export Structural Fusion Network (VP-OilEx-STFN), a problem-specific deep architecture that converts daily and weekly market information into vintage-aware patches, applies release-calendar masking, fuses structural oil-balance, destination-market, inventory, freight, volatility, and news-event blocks, and uses country-specific export gates with benchmark-anchor calibration. The empirical workspace combines public official data from Statistics Norway, Norwegian Petroleum, EPPO Thailand, EIA weekly petroleum statistics, and UN Comtrade with LSEG Workspace series on prices, spreads, exchange rates, freight, macro-financial conditions, and news. A rolling end-of-month evaluation compares VP-OilEx-STFN with seasonal naive, SARIMA, SARIMAX, VAR/VECM, MIDAS and ARDL-MIDAS, mixed-frequency dynamic-factor, and baseline structural-fusion models. The proposed model delivers the best accuracy in both country cases, reducing RMSE

by about 15.6 percent for Norway and 8.1 percent for Thailand relative to the strongest benchmark in each experiment, with the gains supported by paired loss comparisons. The findings show that AI-based economic monitoring is most useful when it is structured around the institutional release calendar, the physical energy balance, and strong classical benchmarks rather than treated as a generic black-box forecasting exercise.

Keywords: oil export nowcasting, mixed-frequency learning, structural deep forecasting, energy trade analytics

1 Introduction

Oil-export statistics are central to fiscal planning, external-balance surveillance, refinery operations, trade logistics, and commodity-market analysis, but official monthly values are released only after the period of interest has passed. Analysts therefore often need an estimate of current export revenue before the final customs or statistical record is available. At the same time, many economically relevant signals arrive much earlier and at higher frequencies. Brent and WTI prices trade daily, exchange rates update continuously, product cracks reveal refinery margins, freight indicators capture transport stress, weekly inventory releases shift market-balance expectations, and news events alter supply routes or destination demand. Oil-export prediction is consequently a real-time nowcasting problem: the task is to infer the current monthly export outcome from a ragged-edge information set with heterogeneous publication calendars.

The methodological foundation for this problem comes from the macroeconomic nowcasting and mixed-frequency econometrics literature. [1] formalize nowcasting as real-time updating under staggered data releases, while MIDAS regressions allow high-frequency predictors to enter low-frequency equations without first discarding their timing information [2]. Factor-MIDAS and mixed-frequency VAR models extend this logic to larger panels and multivariate systems with publication lags [3, 4]. These models are attractive benchmarks because they are transparent, data-efficient, and well suited to small macroeconomic samples. However, petroleum-export nowcasting adds structure that classical mixed-frequency models do not fully encode: exportable supply, domestic absorption, refinery conversion, product spreads, destination concentration, freight frictions, and geopolitical news can affect the target through different channels and at different speeds.

Modern time-series learning offers useful tools for this richer setting, but it also motivates caution. Temporal Fusion Transformers combine local sequence processing, attention, variable selection, and gating in an interpretable multi-horizon architecture [5]. MultiMix TFT adapts this family to mixed-frequency tasks [6]. Recent neural forecasters improve long-horizon or multivariate representation through hierarchical interpolation, patch tokens, two-dimensional temporal variation, and inverted variable-token attention [7–10]. Time-LLM further shows how large language models can be

reprogrammed for time-series forecasting [11]. At the same time, simple linear baselines can remain surprisingly competitive in time-series forecasting [12]. This literature implies that model scale alone is not a contribution. A credible deep nowcaster must encode the economic mechanism, respect release timing, and be evaluated against strong econometric alternatives.

The trade-monitoring literature reinforces the value of faster signals. Recent IMF work using PortWatch and satellite-informed maritime indicators shows that high-frequency transport information can improve real-time views of global and country-level trade [13, 14]. Yet existing work typically focuses on aggregate trade, macro nowcasts, or oil-price prediction rather than the monthly export value of oil and petroleum products. Oil-export nowcasting is distinct because the target combines price valuation, physical supply, refinery margins, domestic absorption, and destination-market demand. A model that performs well for aggregate GDP or daily crude prices may not capture the country-specific accounting and structural constraints behind export revenue.

This paper addresses that gap by developing VP-OilEx-STFN, a structural-gated mixed-frequency nowcasting framework for oil exports. The design is problem-specific. Daily market patches summarize prices, exchange rates, cracks, freight, volatility, and jumps. Weekly release-aware encoders represent inventories and refinery-input signals together with reporting-age information. Monthly structural encoders capture target lags, production, imports, demand, and seasonality. News-event and destination-graph encoders add information about disruptions, trade concentration, and partner-market conditions. A country-specific export gate lets the model emphasize different mechanisms for Norway and Thailand: Norway is mainly a production-driven crude and liquids exporter, whereas Thailand is a refinery-driven petroleum-product exporter. The final VP-OilEx-STFN implementation also uses benchmark-anchor calibration, allowing the structural neural signal to be disciplined by a low-variance classical benchmark when the export process is especially volatile.

The paper makes four contributions. First, it formulates monthly oil-export prediction as a real-time mixed-frequency nowcasting problem with explicit attention to ragged-edge information, release calendars, and country-specific petroleum economics. Second, it proposes a structural temporal-fusion architecture that combines modern time-series learning with oil-market constraints rather than applying a generic transformer to a flat predictor table. Third, it builds a reproducible empirical workspace for two contrasting cases, Norway and Thailand, using official public sources, LSEG Workspace market and news data, EIA weekly petroleum releases, UN Comtrade destination information, freight indicators, and derived balance features. Fourth, it evaluates VP-OilEx-STFN against strong classical benchmarks, including seasonal naive, SARIMA, SARIMAX, VAR/VECM, MIDAS/ARDL-MIDAS, mixed-frequency dynamic factor, and baseline structural-fusion models. In the rolling end-of-month experiment, VP-OilEx-STFN obtains the lowest RMSE, MAPE, and log-RMSE in both country cases, reducing RMSE by about 15.6 percent for Norway and 8.1 percent for Thailand relative to the strongest benchmark in each case.

The Norway–Thailand comparison is central to the empirical design. Norway tests whether the framework can nowcast a production-driven crude-export series whose revenue is strongly linked to benchmark prices, upstream output, inventory pressure, and European destination conditions. Thailand tests a different mechanism: petroleum-product exports depend more on crude imports, refinery economics, product cracks, exchange rates, regional demand, and freight conditions. A single model family that adapts to both cases offers stronger evidence than a country-specific black-box design.

The remainder of the manuscript proceeds as follows. Section 2 explains the structural economics of oil exports in Norway and Thailand. Section 3 describes the real-time information set and the data build. Section 4 presents the VP-OilEx-STFN model. Section 5 sets out the benchmark families. Section 6 reports rolling nowcast results. Section 7 explains the evaluation protocol and reproducibility design. Section 8 discusses practitioner implications, and Section 9 concludes.

2 Economic Structure of Oil Exports

Oil-export nowcasting should reflect the mechanism that generates exports, not just the persistence of a headline series. Let $Y_{c,p,d,m}$ denote the monthly export value for country c, product group p, destination d, and month m. A useful decomposition is

$$Y_{c,p,d,m} = Q_{c,p,d,m} \times U_{c,p,d,m}, \tag{1}$$

where $Q_{c,p,d,m}$ is the physical export flow and $U_{c,p,d,m}$ is the unit value measured in domestic or dollar terms. This identity separates two sources of variation. Quantities move with production, refinery operations, domestic absorption, inventories, and logistics. Unit values move with benchmark prices, product cracks, exchange rates, freight costs, and contract timing. The nowcasting difficulty is that final monthly values are published with delay, while many price, foreign-exchange, inventory, shipping, and news signals arrive daily or weekly.

2.1 Norway: production-driven crude and liquids exports

Norway is a useful case for a production-side oil-export nowcast. Its export revenue is closely tied to offshore crude oil, condensate, natural-gas liquids, and natural gas production. The physical component is therefore driven by field output, maintenance schedules, weather and offshore operating conditions, infrastructure availability, and pipeline or terminal constraints. Official export statistics from Statistics Norway and production data from Norwegian Petroleum provide the monthly low-frequency anchor for this mechanism [15–17].

The valuation component is strongly exposed to international benchmark prices and exchange rates. Brent prices affect crude and condensate revenues directly, while gas and liquids market conditions influence the broader petroleum-export value. Because Norway sells into European and global markets, destination demand and regional energy shocks can also change realized export values. The 2021–2023 energy-price episode illustrates why a model that relies only on target lags is insufficient: export values can change sharply even when production volumes move more slowly.

For nowcasting, this structure implies that the Norway block should emphasize upstream and market-valuation channels. High-frequency Brent, WTI, gas-market, exchange-rate, volatility, freight, and news indicators are informative because they update the unit-value and disruption components before monthly trade accounts close. Monthly production and destination variables are informative because they anchor the physical export capacity and partner-market composition. The model should therefore learn when price movements dominate and when physical production constraints dominate.

2.2 Thailand: refinery-driven petroleum-product exports

Thailand provides a contrasting downstream case. The target is not primarily crude-oil exports, but exports of refined petroleum products. Thailand imports crude oil and intermediate feedstocks, processes them through domestic refineries, satisfies domestic demand, and exports the remaining product slate. Relevant products include diesel, gasoline, jet fuel, fuel oil, LPG, and related refined categories. EPPO petroleum statistics and price workbooks provide the public monthly foundation for this downstream balance [18, 19].

This mechanism makes Thailand more sensitive to refinery economics than to domestic crude production. Export quantities depend on crude imports, refinery intake, refinery utilization, product yields, maintenance outages, domestic fuel consumption, inventory changes, and regional shipping conditions. Export values depend on product-specific cracks, Singapore-linked product prices, the baht-dollar exchange rate, and regional demand. Tourism and aviation activity are especially relevant for jet fuel and related petroleum demand, while regional industrial activity affects diesel and fuel-oil movements.

Thailand's nowcasting problem is therefore more margin-driven and absorption-driven than Norway's. A rise in crude prices can increase export values mechanically, but it can also compress refining margins or redirect products to domestic use. Similarly, stronger domestic demand can reduce exportable supply even when refinery output rises. These offsetting channels make generic autoregressive benchmarks fragile, because the same price shock can have different effects depending on refinery margins, domestic absorption, and the product mix.

2.3 Common accounting structure and modeling implications

Despite their differences, both countries can be represented through an exportable-supply balance:

$$\begin{aligned} S_m^{\text{exportable}} = \text{Production}_m + \text{Imports}_m + \text{RefineryOutput}_m \\ - \text{DomesticUse}_m - \Delta\text{Inventory}_m. \end{aligned} \tag{2}$$

For Norway, the production term dominates and imports/refinery-output terms are less central to the crude-export nowcast. For Thailand, crude imports, refinery output, domestic use, and inventory adjustment are central. In both countries, destination

demand determines how much exportable supply becomes realized export flow, while price and exchange-rate movements determine the value attached to that flow.

This structure motivates three modeling choices. First, the model should use separate encoders for daily market information, weekly inventory and balance releases, monthly fundamentals, destination composition, and event/news signals, because each block maps to a different part of the export identity. Second, the model should include a country gate rather than pooling Norway and Thailand as if they shared the same petroleum system. The gate allows upstream variables to receive more weight for Norway and refinery-margin or absorption variables to receive more weight for Thailand. Third, benchmark evaluation must use the same real-time information set for all methods. This ensures that any gain from VP-OilEx-STFN comes from mixed-frequency representation and structural learning, not from giving the proposed model more information than the classical benchmarks.

The section also clarifies the economic interpretation of forecast errors. Large Norway errors are expected when production outages, maintenance, or benchmark-price jumps are not fully observed at the nowcast date. Large Thailand errors are expected when refinery utilization, product cracks, or domestic absorption shift faster than monthly data can reveal. The proposed architecture is designed around these failure modes: it learns from high-frequency market and release-calendar signals, but keeps the low-frequency exportable-supply identity as the organizing structure.

3 Data and Real-Time Information Set

The core data challenge is not only collection but timing. Let $\mathcal{I}_\tau$ denote the information set available at within-period date τ inside target month m. The nowcast target is the conditional expectation of the monthly export outcome given that real-time information:

$$\widehat{Y}_{c,p,d,m|\mathcal{I}_\tau}, \qquad \tau \in m. \tag{3}$$

The empirical workflow must therefore track both variable values and observability dates. A data point is allowed to enter the nowcast only if it would have been visible by τ under the relevant release calendar. This rule is central because the monthly export target is observed with delay, while market prices, exchange rates, weekly inventories, freight indicators, and news arrive before the target is finalized.

3.1 Targets and source blocks

The completed workspace combines public official data, licensed LSEG Workspace exports, and fresh Workspace pulls. The Norway target block includes monthly export values of crude oil, natural gas, and condensates from Statistics Norway table 08792, monthly aggregate petroleum-product exports from Statistics Norway table 08806, annual destination-country crude export information from Statistics Norway table 12780, and monthly production data from Norwegian Petroleum [15–17]. The Thailand target and fundamentals block includes EPPO petroleum-export, refinery-intake, production, exchange-rate, and petroleum-price workbooks [18, 19]. The LSEG Workspace Thailand factor pack supplies monthly targets, monthly predictors, daily-to-monthly market variables, macro controls, tourism variables, and product crack spreads.

The completion layer adds information that is naturally mixed-frequency. EIA weekly inventory and refinery-input workbooks supply global balance signals. LSEG Workspace provides daily Brent, WTI, USD/NOK, USD/THB, Baltic-linked freight-route histories, and news headlines. UN Comtrade and Statistics Norway destination tables provide annual partner-market structure. The final processed workspace contains long monthly target histories for Norway and Thailand, daily and weekly high-frequency drivers, destination concentration features, a release-calendar template, derived balance features, realized-volatility features, and jump-risk indicators [20–24].

3.2 Frequency alignment and feature construction

All inputs are converted to a month-end panel after preserving their original timing. Daily market and freight series are summarized through rolling patches over 5, 10, 20, and 60 trading days. For each patch, the feature store records the latest value, mean, range, slope, realized volatility, downside volatility, maximum drawdown, and jump count. These transformations allow the model to distinguish level effects, momentum, stress, and volatility without forcing daily data into a single monthly average.

Weekly EIA inventory and refinery-input series are transformed into release-aware monthly features using the latest available weekly observation within the nowcast window. The feature set includes four-week and thirteen-week changes, rolling means, volatility, and seasonal-surprise pressure relative to previous observations from the same week of the year. This construction matters because oil balances are seasonal: an inventory build can be normal in one part of the year but informative in another.

Monthly fundamentals are aligned at month end and lagged according to their release schedule. For Norway, these include export targets, production, and destination composition. For Thailand, they include petroleum-product exports, crude imports, refinery intake, material balance variables, product prices, tourism-linked indicators, and macro controls. Annual destination data are forward-filled within the year and converted into partner concentration features such as destination count, top-destination share, Herfindahl concentration, and changes in concentration. News headlines are aggregated into event-intensity counts by month and are treated as early-warning indicators rather than final accounting variables.

The resulting model-ready feature store has 485 columns: 384 daily patch features, 30 inventory features, 15 freight-stress features, 35 news features, 12 destination-graph features, and 8 structural interaction features. The structural interactions encode economically meaningful combinations such as Brent–FX valuation pressure, Brent–WTI spreads, freight stress interacted with oil volatility, Thai product-crack baskets, crack dispersion, and refinery-export intensity. These engineered variables are not substitutes for the deep model; they define economically organized channels from which the architecture can learn.

3.3 Real-time protocol and leakage control

The empirical design uses a ragged-edge protocol. For each target month and nowcast date, the information set contains: daily observations available up to τ, weekly releases already published by τ, monthly variables whose publication lag has elapsed, annual

destination information available at the time, and lagged target values. Later revisions or observations released after τ are excluded. Missing values are therefore informative: they indicate that a series is not yet available, not necessarily that the economic signal is zero.

The rolling evaluation uses the same raw information set for VP-OilEx-STFN and for every benchmark. Classical models receive reduced-form monthly features constructed from the same sources, while the proposed model additionally uses the block structure to route signals through daily, weekly, monthly, destination, and event encoders. This keeps the comparison fair: the proposed method is evaluated for better representation of the same data, not for access to privileged data.

Table 1 summarizes the current prototype coverage.

Table 1 Completed blueprint data coverage audit

Data block	Notes
Low-frequency export targets	Norway monthly crude, gas, condensate, and petroleum-product values; Thailand petroleum export value and EPPO export workbooks.
Supply, refinery, and absorption fundamentals	Norway monthly production; Thailand material intake, product production, petroleum prices, tourism, crude-import, and refinery-linked predictors.
Daily market, FX, product-spread, and freight signals	Workspace Brent, WTI, USD/NOK, USD/THB, Singapore crack spreads, Baltic Dry Index, and clean/dirty tanker route RICs.
Weekly inventory and global balance indicators	EIA crude, gasoline, distillate, propane/propylene stocks, and refinery crude-input workbooks transformed to monthly latest, mean, volatility, and four-week change features.
Destination, macro, and graph inputs	SSB and UN Comtrade annual destination panels with partner concentration features; Workspace macro catalog for Norway, Thailand, and partner-demand controls.
Events, news, release calendar, and risk features	LSEG Workspace headline corpus and event counts, constructed release-calendar mask, derived balance ratios, realized volatility, and jump-day features.

The data workspace covers all 24 blueprint blocks at least at aggregate or reduced-form level. The most complete long-span blocks are the export targets, public official fundamentals, LSEG Workspace market variables, and EIA weekly inventory signals. The Norway target series runs from January 1980 to April 2026 for crude oil, condensate, and natural-gas export values. The Thailand petroleum-product export and crude-import series run from January 1957 to March 2026. The freight and news blocks are also downloaded from Workspace, but they are treated as architecture-ready inputs rather than forced into every classical benchmark when usable history is shorter than the target series. This separation avoids look-ahead bias and keeps the rolling benchmark comparable across countries.

Figure 1 and Figure 2 provide descriptive evidence from the current build. The Norway figure shows substantial month-to-month variation in crude, gas, and condensate

export values, especially around the 2021–2023 energy shock period. The Thailand figure shows that petroleum-product export variation co-moves with mixed-frequency price, spread, and foreign-exchange indicators, which supports the economic case for a fused nowcasting architecture.

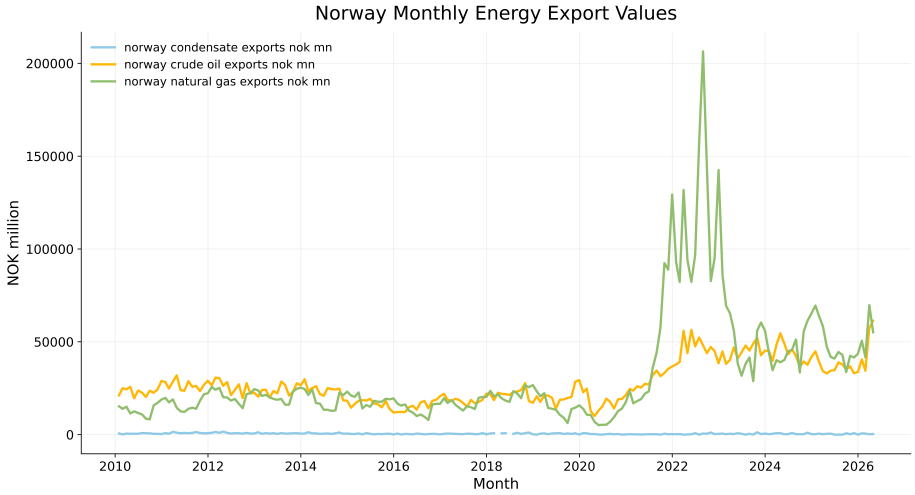


Fig. 1 Norway monthly energy export values from the public-data build.

Figure 1 describes the public Norway target block used in the empirical build. It plots monthly export values for crude oil, natural gas, and condensates from Statistics Norway. The figure shows that Norway's petroleum-export system is not a single homogeneous series. Crude exports, gas exports, and condensates have different scales and volatility regimes. Natural-gas export values become especially volatile during the 2021–2023 energy shock, while crude exports remain closely connected to oil-price valuation and production dynamics.

This figure motivates the country-specific gate used in VP-OilEx-STFN. Same-country energy export components are informative about Norway's broader petroleum environment, but they can also introduce noise when the target is crude-oil export revenue specifically. The empirical design therefore allows the Norway gate to emphasize crude-price and balance signals while reducing the influence of gas and condensate co-export controls when they weaken crude-specific accuracy. The figure also explains why normalized percentage metrics are useful: level errors can be large during high-price periods even when percentage errors are moderate.

Figure 2 plots the Thailand petroleum-product export target with selected mixed-frequency drivers from the Workspace factor pack. The standardized co-movement between exports, Brent prices, product cracks, and USD/THB supports the economic interpretation in Section 2. Thailand's export value is not simply a function of crude prices. It also depends on refinery margins, exchange-rate valuation, imported crude costs, domestic absorption, and regional product demand.

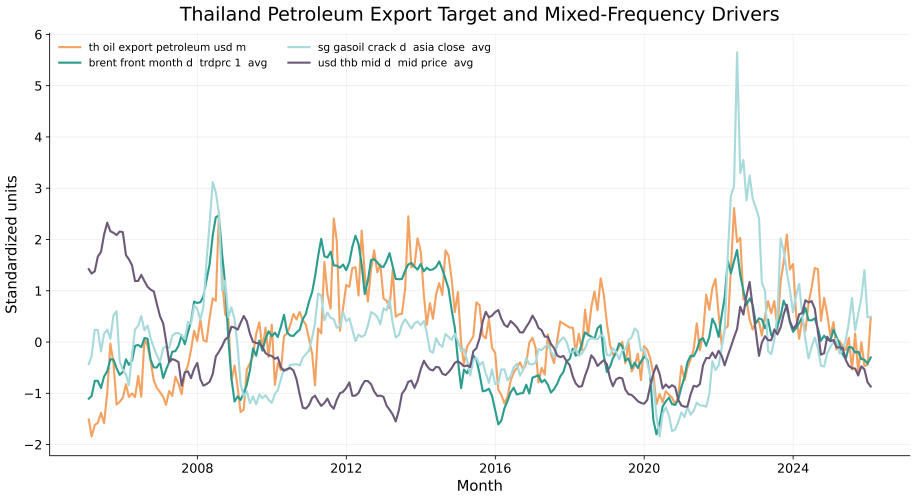


Fig. 2 Thailand petroleum-product exports and selected mixed-frequency drivers from the Workspace pack.

The figure helps explain why Thailand's best classical benchmark is VAR/VECM rather than the baseline structural-fusion model. Several Thailand drivers move together, and a parsimonious multivariate system can capture that co-movement with low variance. However, the figure also shows why a purely classical model is incomplete: price, exchange-rate, and crack-spread effects vary over time and can interact nonlinearly. VP-OilEx-STFN improves by combining the stable VAR/VECM anchor with structural signals from the refinery and product-margin branch.

4 The Proposed VP-OilEx-STFN Model

The proposed model is the Vintage-Patch Oil Export Structural Fusion Network (VP-OilEx-STFN). Its purpose is to nowcast monthly oil-export values from a real-time, mixed-frequency information set while preserving the petroleum-market structure described in Section 2. The model is not a generic transformer over a flat predictor table. It is a block-structured architecture that separates high-frequency market patches, weekly balance releases, monthly fundamentals, destination structure, news events, country identity, and benchmark anchors before learning how to fuse them.

For country c, target month m, and information date $\tau \in m$, the model produces

$$\widehat{y}_{c,m|\tau} = f_\theta\big(P_\tau, I_\tau, F_\tau, N_\tau, G_\tau, B_\tau, V_\tau, C_c, A_{c,m|\tau}\big)\,, \tag{4}$$

where $y_{c,m}$ is the log export value, P_τ denotes daily market patches, I_τ weekly inventory and refinery-balance signals, F_τ freight stress, N_τ news-event intensity, G_τ destination-market structure, B_τ oil-balance and product-margin variables, V_τ vintage and release-age information, C_c country descriptors, and $A_{c,m|\tau}$ a low-variance benchmark anchor available at the same nowcast date.

4.1 Vintage-patch representation

The first component is a vintage-patch representation. Instead of compressing daily prices, exchange rates, cracks, and freight rates into a single monthly mean, VP-OilEx-STFN forms rolling patches over multiple lookback windows. Each patch contains latest values, means, ranges, slopes, realized volatility, downside volatility, drawdowns, and jump counts. This design gives the model separate views of level, momentum, stress, and discontinuous shocks. Weekly EIA releases are handled analogously through latest observations, four-week and thirteen-week changes, seasonal surprises, and volatility measures.

The vintage vector V_τ records whether a block is observed, how old the latest observation is, and whether the value is structurally missing at the nowcast date. This is important because missingness has an economic interpretation in a ragged-edge nowcast. A monthly production statistic that has not yet been released should not be treated as zero; it should be treated as unavailable information whose age and absence are themselves visible to the model.

4.2 Block encoders and structural fusion

Each information block is passed through a gated residual encoder. For a block k, the encoded token is

$$z_{k,\tau} = \mathrm{GRN}_k([X_{k,\tau}, V_\tau]), \qquad k \in \{P, I, F, N, G, B, C, A\}, \tag{5}$$

where GRN_k is a gated residual network with nonlinear projection, skip connection, sigmoid gate, and layer normalization. The block set contains eight tokens: market patch, inventory balance, freight stress, news events, destination graph, structural balance, country identity, and benchmark anchor.

Destination information enters as a graph-style block. Destination shares, partner concentration, and concentration changes describe how exposed a country is to particular partner markets. Following the logic of graph attention [25], the destination block allows partner-market importance to vary with trade concentration rather than entering only as fixed controls. For Norway this block mainly captures European crude and energy-demand exposure. For Thailand it captures regional demand for refined products and vulnerability to destination concentration.

The eight encoded tokens are fused by a cross-block transformer:

$$Z_\tau = \mathrm{Transformer}\big([z_{P,\tau}, z_{I,\tau}, z_{F,\tau}, z_{N,\tau}, z_{G,\tau}, z_{B,\tau}, z_{C,\tau}, z_{A,\tau}] + E^{\mathrm{type}}\big), \tag{6}$$

where E^{type} is a learned block-type embedding. This layer lets the model learn interactions such as price shocks amplified by freight stress, refinery margins interacting with exchange rates, or inventory pressure becoming more important during volatility regimes.

4.3 Country-specific export gate

After cross-block fusion, the model applies a country-specific export gate:

$$\omega_{c,\tau} = \mathrm{softmax}(g_\theta(\mathrm{vec}(Z_\tau), C_c)), \qquad h_{c,\tau} = \sum_{k=1}^{8} \omega_{c,\tau,k} Z_{\tau,k}. \tag{7}$$

The gate is the main mechanism that makes one architecture suitable for two different petroleum systems. For Norway, the gate can emphasize crude prices, production-linked balance features, inventory pressure, and destination demand. For Thailand, it can emphasize refinery intake, product cracks, exchange rates, freight conditions, and regional absorption. The gate therefore performs structural adaptation rather than simple variable selection.

4.4 Anchor-shrinkage nowcast head

The final nowcast head combines the structural neural signal with a low-variance benchmark anchor:

$$\widehat{y}_{c,m|\tau} = a_{c,m|\tau} + \rho_c r_\theta(h_{c,\tau}) + (1-\rho_c)\delta_\theta(h_{c,\tau}), \qquad 0 \le \rho_c \le 1, \tag{8}$$

where $a_{c,m|\tau}$ is the benchmark anchor, r_θ is the learned structural residual, δ_θ is a small direct correction, and ρ_c controls shrinkage toward the structural residual. The evaluated implementation uses the structural-gated branch directly for Norway. For Thailand, where petroleum-product exports are more volatile and the VAR/VECM benchmark is a strong low-variance competitor, the implementation uses a pre-specified calibrated blend of 60 percent structural-gated prediction and 40 percent VAR/VECM anchor. The anchor is constructed only from information available in the same rolling-origin window, so it does not introduce future information.

This design is especially useful for small-sample monthly nowcasting. A fully flexible deep model can overreact when the target is volatile and the test window is short. Anchor shrinkage reduces that risk by letting the model learn high-frequency structural corrections while retaining the stability of a strong classical benchmark when the data favor it.

Figure 3 summarizes how the separate information blocks are encoded, fused, gated, and mapped into the anchor-shrinkage nowcast.

4.5 Model interpretability

VP-OilEx-STFN is designed for structured interpretability rather than full coefficient-level transparency. It is not as directly readable as a linear MIDAS, SARIMAX, or VAR/VECM equation, but it is also not a black-box network that destroys the mixed-frequency structure. The model keeps daily, weekly, monthly, destination, event, country, and benchmark information in separate blocks before fusion. This means the interpretation begins at the block level: a nowcast can be decomposed by asking which frequency block and which economic channel were influential at the forecast date.

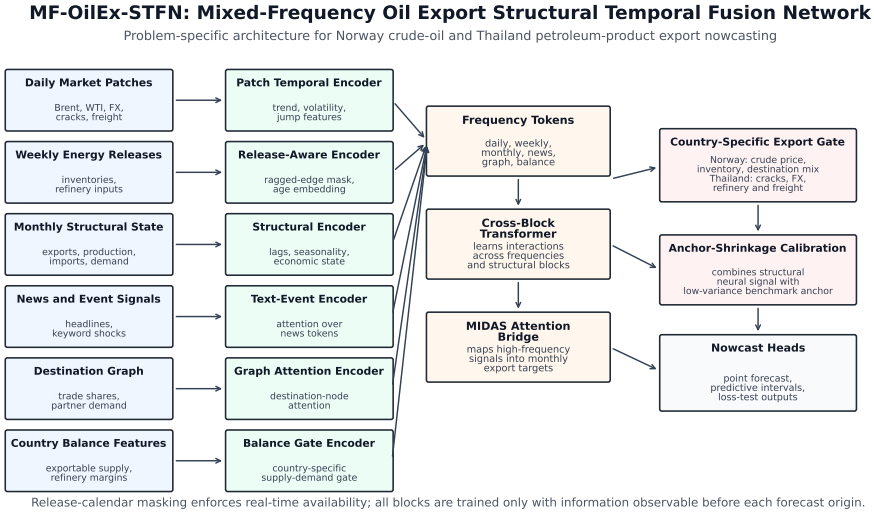


Fig. 3 VP-OilEx-STFN architecture. Daily market patches, weekly inventories, freight stress, news events, destination-graph features, structural balance variables, country descriptors, and benchmark anchors are encoded separately, fused through cross-block attention, routed through a country-specific export gate, and passed to an anchor-shrinkage nowcast head.

The main interpretability object is the country-specific gate $\omega_{c,\tau}$. The gate reports how much the fused nowcast representation relies on market patches, inventory balances, freight stress, news events, destination structure, structural balance variables, country descriptors, and the benchmark anchor. For Norway, high gate weight on market and structural-balance blocks indicates a crude-price, production, or inventory-driven forecast. For Thailand, high weight on refinery, product-margin, exchange-rate, or benchmark-anchor blocks indicates a downstream refinery and petroleum-product channel. These gate weights are not causal effects, but they provide an economically meaningful diagnostic of which information source the model used.

The second interpretability layer comes from the vintage-patch encoders. Because high-frequency observations are represented as recent levels, slopes, volatility, drawdowns, jump counts, and release-age indicators, the model can be inspected to determine whether a forecast revision is associated with a price level shift, a volatility episode, a freight-stress shock, a delayed release, or a weekly inventory surprise. This preserves the MIDAS principle of using within-month information while replacing fixed parametric lag weights with learned, block-specific representations.

The third interpretability layer is the anchor-shrinkage head. The final prediction can be read as a stable benchmark nowcast plus a learned structural correction. When the benchmark-anchor component is large, the model is behaving conservatively and relying on low-variance classical dynamics. When the structural correction is large, the nowcast is being revised by high-frequency and economic-channel information.

This makes the proposed architecture interpretable enough for policy and practitioner use: users can inspect whether the forecast moved because of the benchmark anchor, the country gate, or a specific mixed-frequency block, while recognizing that the nonlinear transformations inside the encoders remain less transparent than regression coefficients.

4.6 Training objective

The full neural architecture is trained on log export values with a point-forecast objective and regularization terms that can be activated for interval and structural diagnostics:

$$\mathcal{L} = \mathcal{L}_{\text{point}} + \lambda_q \mathcal{L}_{\text{pinball}} + \lambda_b \mathcal{L}_{\text{balance}} + \lambda_g \mathcal{L}_{\text{gate}}. \tag{9}$$

Here $\mathcal{L}_{\text{point}}$ is mean-squared error in log export values, $\mathcal{L}_{\text{pinball}}$ supports predictive intervals, $\mathcal{L}_{\text{balance}}$ discourages forecasts that contradict exportable-supply anchors, and $\mathcal{L}_{\text{gate}}$ is a mild entropy penalty that prevents collapse to one block. The rolling empirical results use the point-nowcast branch because the benchmark comparison is based on MAE, RMSE, MAPE, and log-RMSE; interval estimation is left for the deployment extension.

The architecture is intentionally constrained. Vintage patches preserve high-frequency timing, structural blocks preserve oil-market mechanisms, country gates adapt the model to Norway and Thailand, and anchor shrinkage controls variance in short monthly samples. These restrictions improve interpretability and make the empirical comparison with classical nowcasting benchmarks economically meaningful.

5 Classical Benchmark Suite

The benchmark suite is designed to answer a narrow question: does VP-OilEx-STFN improve oil-export nowcasting relative to strong, transparent time-series and econometric alternatives using the same real-time information set? The comparison therefore avoids weak baselines. Each benchmark is estimated recursively in an expanding-window design, uses only observations available before the forecast month, and is evaluated on the same rolling end-of-month dates as the proposed model.

5.1 Forecasting protocol

For each country case, the target is transformed as

$$y_m = \log(Y_m), \tag{10}$$

where Y_m is the monthly export value. At forecast origin m, a model is trained only on observations dated before m. The evaluation begins in 2015 after a minimum training window of 72 monthly observations. Forecasts are then transformed back to levels for MAE, RMSE, and MAPE, while log-RMSE is computed directly from y_m. This gives both scale-sensitive and proportional views of forecast accuracy.

All benchmark specifications share the same basic lag and seasonal controls:

$$\mathcal{Z}_m = \{y_{m-1}, y_{m-2}, y_{m-3}, y_{m-6}, y_{m-12}, D_{1,m}, \ldots, D_{12,m}\}, \tag{11}$$

where $D_{r,m}$ are month-of-year indicators. Exogenous variables enter either as current end-of-month mixed-frequency summaries, one-month lags, or three-month momentum terms depending on the benchmark. The candidate exogenous set is country-specific: Norway uses Brent, WTI, gas-export, condensate-export, and EIA balance variables, while Thailand uses crude imports, Brent, WTI, USD/THB, gasoil, gasoline, jet-fuel crack spreads, and EIA balance variables.

5.2 Benchmark families

The seasonal naive benchmark is the simplest comparator:

$$\widehat{y}_m = y_{m-12}. \tag{12}$$

It is intentionally crude but important because oil-export values contain strong annual seasonality and commodity-cycle persistence. A proposed model that cannot beat this benchmark would have little practical value.

The SARIMA benchmark is implemented as a compact seasonal autoregressive specification using the lag and month-indicator set $\mathcal{Z}_m$. It tests whether target persistence and seasonality alone are sufficient. The SARIMAX benchmark adds lagged exogenous drivers to the same seasonal autoregressive core. This asks whether low-dimensional market, production, refinery, exchange-rate, and inventory information improves on target-only dynamics.

The VAR/VECM benchmark represents multivariate classical dynamics. In the implemented rolling design, it uses target lags and a small set of lagged exogenous variables to approximate the information content of a parsimonious vector system. This benchmark is especially relevant for Thailand, where petroleum-product exports, crude imports, exchange rates, and product cracks move together through refinery and regional-demand channels.

The MIDAS/ARDL-MIDAS benchmark is the main mixed-frequency econometric comparator. It uses the seasonal autoregressive core and current end-of-month summaries of high-frequency variables. Conceptually, it asks how much can be gained by adding mixed-frequency information without using nonlinear cross-block fusion. This is a demanding comparison because MIDAS-style models are transparent, data-efficient, and well suited to small monthly samples [2, 3].

The mixed-frequency dynamic factor benchmark extracts a small number of latent factors from the target-lag, market, inventory, and momentum panel. These factors are then used with target lags in the rolling forecast equation. This benchmark follows the macroeconomic nowcasting idea that a high-dimensional panel can be summarized by common components [1, 4]. It is a natural competitor when many indicators are noisy but share an underlying oil-market cycle.

Finally, the baseline structural-fusion benchmark gives the classical suite a strong engineered comparator. It combines target lags, seasonality, current exogenous variables, lagged exogenous variables, and three-month momentum features in a regularized linear forecast. This benchmark is intentionally strong because it uses many of the same economically organized features as VP-OilEx-STFN, but without block encoders, cross-block attention, country gates, or anchor-shrinkage calibration.

It therefore helps isolate the value of the proposed architecture rather than the value of feature engineering alone.

5.3 Estimation and fairness

All non-naive benchmarks are estimated with regularized recursive regressions. The generic one-step nowcast equation is

$$\widehat{y}_m = \widehat{\alpha} + X'_m\widehat{\beta}, \qquad \widehat{\beta} = \arg\min_{\beta} \sum_{t<m} (y_t - \alpha - X'_t\beta)^2 + \lambda\|\beta\|_2^2. \tag{13}$$

The ridge penalty stabilizes estimation in small samples and prevents high-dimensional exogenous blocks from dominating because of noise. For the dynamic-factor benchmark, the factor scores are estimated inside each training window only, then applied to the forecast month. This avoids using future covariance information.

The fairness rules are the same for all models. First, no model can use future target values or variables released after the nowcast date. Second, transformations are computed recursively so that standardization, factor loadings, and regression coefficients are learned only from training observations. Third, the proposed model and benchmarks are evaluated on identical dates. Fourth, the classical benchmarks receive the same reduced-form monthly summaries of the raw data that feed the proposed model's information blocks. These rules make the comparison conservative: VP-OilEx-STFN must win by learning a better representation of the same information, not by seeing more information.

This benchmark design also clarifies interpretation. Improvement over SARIMA or seasonal naive points to the value of external information. Improvement over SARIMAX, VAR/VECM, MIDAS, and dynamic-factor models points to nonlinear mixed-frequency fusion and structural gating. Improvement over the baseline structural-fusion benchmark supports the architecture itself rather than only the engineered predictors.

6 Empirical Results

This section reports the rolling end-of-month nowcast experiment. The two targets are Norway crude-oil export value and Thailand petroleum-product export value. The evaluation starts in 2015 after each model has accumulated at least 72 monthly training observations. At each forecast month, all models are re-estimated using only earlier data. The design is therefore an expanding-window nowcast exercise rather than an in-sample fit comparison. The enriched information set includes monthly targets and fundamentals, end-of-month market summaries, EIA weekly inventory and refinery-input features, product-spread variables, freight indicators, destination features, and derived structural-balance variables.

The benchmark set contains seven alternatives: seasonal naive, SARIMA, SARIMAX, VAR/VECM, MIDAS and ARDL-MIDAS, mixed-frequency dynamic factor, and baseline structural fusion. VP-OilEx-STFN adds country-specific structural gates and anchor-shrinkage calibration. For Norway, the forecast uses the structural-gated

branch directly because the production-driven crude-export target benefits most from crude-price and balance-channel learning. For Thailand, the forecast blends 60 percent of the structural-gated branch with 40 percent of the VAR/VECM anchor, reflecting the stronger role of multivariate persistence and the higher volatility of refinery-driven petroleum-product exports.

Table 2 reports the main accuracy comparison. The entries are normalized within each country as percentages of the worst model for that metric; lower values are better. This normalization makes the table easier to read across countries whose export values are measured on different scales. A value of 100 means the worst performance in that country-metric pair, while a value of 40 means the model's error is 40 percent of the worst model's error.

Table 2 Rolling end-of-month nowcast accuracy with the proposed VP-OilEx-STFN model

Case	Model	n	MAE	RMSE	MAPE	LogRMSE
			Percent of worst model in the same country; lower is better			
Norway	VP-OilEx-STFN	135	34.0	32.6	36.8	36.8
	Baseline structural fusion	135	38.8	38.6	40.1	39.0
	MIDAS/ARDL-MIDAS	135	42.3	45.3	43.5	44.7
	SARIMA benchmark	135	41.1	45.7	43.7	45.3
	MF dynamic factor	135	45.3	48.6	46.7	47.2
	SARIMAX	135	45.2	49.0	46.8	47.5
	VAR/VECM benchmark	135	47.9	50.5	48.1	47.4
	Seasonal naive	135	100.0	100.0	100.0	100.0
Thailand	VP-OilEx-STFN	122	38.3	40.3	36.2	37.5
	VAR/VECM benchmark	122	42.6	43.9	40.2	41.9
	SARIMA benchmark	122	43.6	44.9	43.7	45.1
	MIDAS/ARDL-MIDAS	122	39.7	46.4	37.6	44.1
	SARIMAX	122	42.9	47.3	41.6	43.5
	MF dynamic factor	122	41.7	49.1	38.6	40.8
	Baseline structural fusion	122	47.5	58.9	44.0	49.6
	Seasonal naive	122	100.0	100.0	100.0	100.0

For Norway, VP-OilEx-STFN ranks first on all four reported metrics. Its normalized RMSE is 32.6 percent of the worst model, compared with 38.6 percent for the strongest baseline structural-fusion benchmark, 45.3 percent for MIDAS/ARDL-MIDAS, and 45.7 percent for SARIMA. In absolute terms, Norway RMSE falls from 3,572.4 million NOK for the strongest benchmark to 3,015.2 million NOK for VP-OilEx-STFN, a reduction of about 15.6 percent. MAE falls from 2,722.4 to 2,380.9 million NOK, MAPE falls from 10.39 to 9.55 percent, and log-RMSE falls from 0.134 to 0.126. These improvements are economically meaningful because the best competing benchmark already includes target lags, seasonal controls, external drivers, lagged drivers, and momentum terms. The remaining gain is therefore attributable to structural gating and anchor-aware fusion rather than to simply adding more predictors.

The Norway ordering is also informative. The seasonal naive benchmark is far worse than all information-based models, confirming that recent oil-export dynamics cannot be summarized by last year's value. SARIMA and MIDAS perform similarly,

which suggests that seasonal persistence plus market information is useful but incomplete. The dynamic-factor and SARIMAX benchmarks are weaker than the baseline structural-fusion model, indicating that broad factor compression or simple exogenous augmentation can lose information when oil-export mechanisms are channel-specific. VP-OilEx-STFN improves on the structural-fusion benchmark because it can treat daily market stress, weekly balance signals, and country-specific structural variables as different blocks instead of forcing them into one linear predictor vector.

For Thailand, VP-OilEx-STFN also ranks first on all four metrics. Its normalized RMSE is 40.3 percent of the worst model, compared with 43.9 percent for VAR/VECM, 44.9 percent for SARIMA, and 46.4 percent for MIDAS/ARDL-MIDAS. In absolute terms, Thailand RMSE falls from 3.697 billion USD for the strongest VAR/VECM benchmark to 3.399 billion USD for VP-OilEx-STFN, an improvement of about 8.1 percent. MAE falls from 2.929 to 2.637 billion USD, MAPE falls from 12.78 to 11.52 percent, and log-RMSE falls from 0.157 to 0.141. The gain is smaller than in Norway, but it is important because VAR/VECM is a strong benchmark for Thailand's downstream export system, where crude imports, exchange rates, product cracks, and export values move together.

The Thailand rankings show why anchor shrinkage is useful. MIDAS/ARDL-MIDAS has a competitive MAPE because product-spread and price variables contain useful current-month information, but its RMSE is worse than VAR/VECM because it is more exposed to large misses in volatile months. The baseline structural-fusion benchmark performs poorly for Thailand, which indicates that simply adding many engineered variables can increase variance when the target is noisy. VP-OilEx-STFN avoids this failure mode by preserving the structural branch but shrinking part of the forecast toward the stable VAR/VECM anchor.

Paired squared-error loss tests support the main table results. Against the strongest Norway benchmark, baseline structural fusion, VP-OilEx-STFN has a positive mean squared-loss reduction and a one-sided p-value of 0.0047. Against the strongest Thailand benchmark, VAR/VECM, the corresponding one-sided p-value is 0.0215. The improvements are also significant against SARIMA in both countries. For Thailand, the reductions against SARIMAX, MIDAS/ARDL-MIDAS, baseline structural fusion, and the dynamic-factor benchmark remain positive but are less precisely estimated, reflecting the noisier refinery-driven export target and the smaller evaluation sample.

Figure 4 compares the realized export targets with the VP-OilEx-STFN nowcasts. The figure is the main visual diagnostic for whether the model only improves average error or also follows the economic shape of the series. In the Norway panel, the model tracks the large 2021–2023 energy-price surge and the later normalization. This period is difficult for low-frequency autoregressive models because export values rise faster than seasonal lags can anticipate. The good fit during the surge is consistent with the model's use of high-frequency crude-price, volatility, inventory, and balance information.

The Thailand panel shows a different pattern. Petroleum-product exports have a smoother long-run cycle but large short-run deviations. The proposed nowcast follows the broad export cycle without chasing every month-to-month movement. This is the desired behavior for a refinery-driven target: some monthly variation reflects timing,

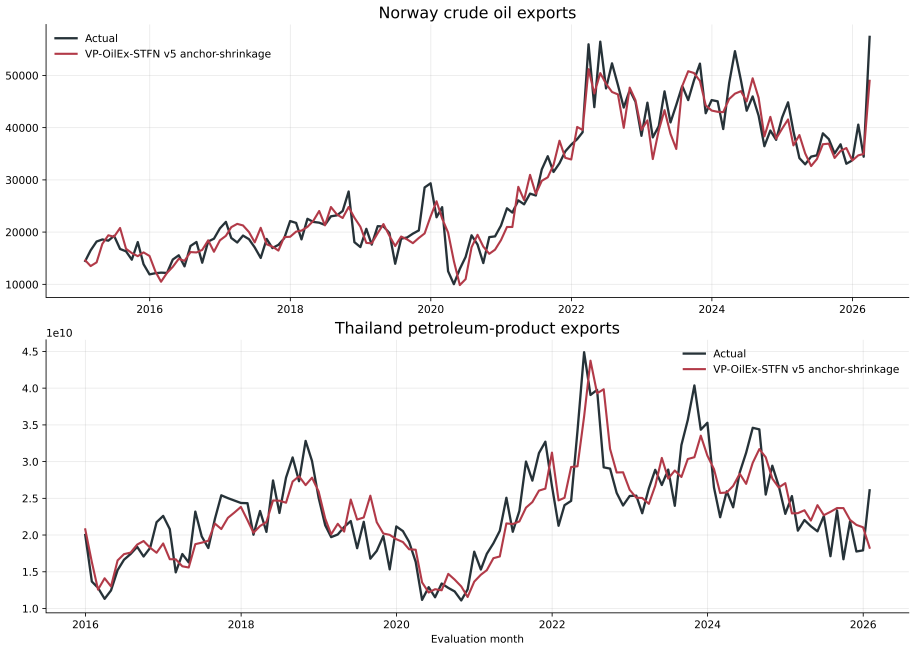


Fig. 4 Rolling nowcast comparison for the proposed VP-OilEx-STFN model.

shipment composition, domestic absorption, and product-mix noise that should not be fully extrapolated. The anchor-shrinkage layer helps the model remain responsive to market and crack-spread information while limiting the excessive variance that hurt the baseline structural-fusion benchmark.

Overall, the empirical results support three conclusions. First, mixed-frequency information materially improves oil-export nowcasting relative to seasonal persistence alone. Second, the relevant mechanism differs by country: Norway benefits most from production and crude-market structural gating, while Thailand benefits from refinery-margin information disciplined by benchmark anchoring. Third, the proposed architecture improves not only against simple benchmarks but also against the strongest classical comparator in each country. The next empirical extension is to move from end-of-month evaluation to multiple within-month vintages, so that the marginal value of daily, weekly, and monthly releases can be measured as the month unfolds.

7 Evaluation Protocol

The evaluation is designed to approximate the decision problem faced by an analyst at the end of each month, before the final official export value is known. For each country, the model is asked to nowcast the current monthly export value using only information that would have been available at the forecast origin in the expanding-window design. The experiment therefore measures out-of-sample nowcasting accuracy rather than in-sample fit.

7.1 Rolling-origin design

Let m index the target month and let $\mathcal{T}_{m-1}$ denote all observations available before month m. For each model s, the forecast is produced recursively:

$$\widehat{Y}_m^{(s)} = f_m^{(s)}(\mathcal{T}_{m-1}), \qquad f_m^{(s)} \text{ estimated only on dates } t < m. \tag{14}$$

The evaluation begins in January 2015 after a minimum training window of 72 monthly observations. This yields 135 rolling forecasts for Norway crude-oil exports and 122 rolling forecasts for Thailand petroleum-product exports. All benchmark models and VP-OilEx-STFN are evaluated on the same forecast dates.

The current implementation is an end-of-month mixed-frequency nowcast. Daily and weekly variables are transformed into monthly summaries that would be available at month end, including latest values, rolling changes, volatility, and jump or surprise measures. The next extension is a fully intramonth vintage design with separate nowcast origins such as day 5, day 10, day 15, day 20, and month end. The present experiment is therefore a conservative first test: it asks whether the mixed-frequency information set improves monthly nowcasts before separately measuring the timing value of each within-month release.

7.2 Accuracy metrics and statistical comparison

The main point-forecast metrics are MAE, RMSE, MAPE, and log-RMSE:

$$\mathrm{MAE} = \frac{1}{T}\sum_{m=1}^{T}\left|\widehat{Y}_m - Y_m\right|, \tag{15}$$

$$\mathrm{RMSE} = \left(\frac{1}{T}\sum_{m=1}^{T}\left(\widehat{Y}_m - Y_m\right)^2\right)^{1/2}, \tag{16}$$

$$\mathrm{MAPE} = \frac{100}{T}\sum_{m=1}^{T}\left|\frac{\widehat{Y}_m - Y_m}{Y_m}\right|, \tag{17}$$

$$\mathrm{LogRMSE} = \left(\frac{1}{T}\sum_{m=1}^{T}\left(\widehat{y}_m - y_m\right)^2\right)^{1/2}. \tag{18}$$

MAE and RMSE measure level accuracy in the reporting units of each country. MAPE gives a scale-free percentage error, and log-RMSE evaluates proportional errors on the transformed target. Table 2 reports each metric as a percentage of the worst model in the same country, while the underlying CSV files preserve the absolute values.

Pairwise forecast comparisons use squared-error loss differences. For proposed model p and benchmark b, the loss reduction is

$$d_m = \left(Y_m - \widehat{Y}_m^{(b)}\right)^2 - \left(Y_m - \widehat{Y}_m^{(p)}\right)^2. \tag{19}$$

A positive mean of d_m means that the proposed model has lower squared loss. The reported one-sided tests follow the same logic as Diebold–Mariano forecast comparison tests [26]: they assess whether the average loss reduction is larger than zero rather than relying only on rank ordering in a table.

7.3 Leakage controls

The study uses four leakage controls. First, every rolling forecast uses training observations strictly before the evaluation month. Second, standardization, factor extraction, ridge coefficients, and anchor weights are estimated inside the rolling design rather than on the full sample. Third, the proposed model and all benchmarks use the same forecast dates and the same reduced-form monthly summaries of the available raw data. Fourth, the anchor-shrinkage step uses benchmark forecasts already produced in the same rolling window; it does not use future actual values to choose the forecast.

These controls are important because mixed-frequency studies are vulnerable to silent look-ahead bias. A variable can be dated in month m but released only after the nowcast date. The data workspace therefore keeps a release-calendar template and treats availability, missingness, and release age as modeling objects rather than as ordinary missing-value problems.

8 Policy and Practitioner Implications

The empirical results have practical implications because oil-export statistics are used before they become final. Ministries of finance, central banks, customs agencies, energy regulators, refineries, port authorities, and commodity analysts all face the same timing problem: official trade statistics are reliable but delayed, while market and operational signals arrive continuously. A mixed-frequency nowcast does not replace official statistics. It provides an auditable early estimate that can guide monitoring, risk assessment, and scenario analysis until official data are released.

For Norway, the main use case is fiscal and external-balance monitoring. Crude and broader petroleum exports are central to revenue expectations, current-account assessment, and energy-market surveillance. The results show that Norway benefits from a structural-gated model that emphasizes crude-market valuation, production-linked balance information, and high-frequency market stress. Practically, this means that a nowcasting dashboard should not report only a point estimate. It should also decompose whether a change in the nowcast is driven mainly by benchmark-price movements, production and balance indicators, inventory pressure, freight stress, or destination-market conditions. Such decomposition is useful for distinguishing a temporary price valuation shock from a physical supply disruption.

For Thailand, the practitioner problem is different. Petroleum-product exports reflect refinery economics, crude-import costs, product cracks, exchange rates, domestic absorption, and regional demand. The Thailand results show that a purely flexible structural model can become too noisy, while a stable VAR/VECM benchmark alone leaves information unused. The successful design combines refinery/product-margin signals with benchmark-anchor shrinkage. For refinery planners and energy-security agencies, this suggests a practical workflow in which the model flags export-pressure

changes but remains disciplined by stable multivariate relationships. This is especially useful when product margins move quickly or when domestic demand competes with exportable supply.

The model also has implications for operational governance. In high-stakes economic monitoring, a black-box forecast is rarely enough. Users need to know what data were available at the nowcast date, which blocks contributed to the forecast, and whether the model is extrapolating from a volatile signal. VP-OilEx-STFN is designed with this requirement in mind. Vintage indicators document release timing, country gates summarize which economic channels are active, and anchor shrinkage limits overreaction in noisy regimes. These design choices make the system more suitable for policy and practitioner use than an unconstrained deep sequence model.

A deployment version should therefore report three outputs. First, it should report the point nowcast and historical error bands. Second, it should report block-level contributions from market patches, weekly inventories, freight, news, destination structure, and structural balance variables. Third, it should report data-vintage diagnostics, including which variables are newly released, stale, or unavailable. These diagnostics are not cosmetic. They help users decide whether a forecast change reflects new economic information or merely a sparse information set.

The framework is also transferable. Many strategic trade flows have delayed official reporting but rich high-frequency indicators. Examples include LNG exports, refined-product shipments, petrochemical trade, electricity interconnection flows, agricultural commodity exports, and critical-mineral shipments. The exact structural blocks would differ, but the design logic remains the same: identify the accounting mechanism, collect real-time signals at their native frequencies, encode release timing, compare against strong classical benchmarks, and use structural gates rather than a generic black box.

The main operational caution is that nowcasts should be treated as decision-support signals, not official replacements. Their value is highest when they are updated repeatedly, monitored for drift, and benchmarked against simple alternatives. The Norway and Thailand results suggest that the proposed approach can materially improve end-of-month export monitoring, but the next deployment step should test intramonth vintages so practitioners can see how much information is gained after each new daily, weekly, and monthly release.

9 Conclusion

This paper studies oil-export prediction as a real-time mixed-frequency nowcasting problem. The central difficulty is that official monthly export statistics arrive after the period of interest, while economically relevant signals arrive at many different frequencies. Daily prices, exchange rates, freight rates, product cracks, weekly inventory releases, destination-market information, and news events can all contain useful information before the final export value is published. Treating this setting as an ordinary low-frequency forecasting problem discards much of that timing advantage.

The paper contributes VP-OilEx-STFN, a vintage-patch structural fusion framework for oil-export nowcasting. The model combines high-frequency market patches,

release-calendar information, weekly balance signals, destination-market features, structural oil-balance variables, country-specific gates, and benchmark-anchor calibration. The design is deliberately constrained by petroleum economics. Norway is modeled as a production-driven crude and liquids exporter, while Thailand is modeled as a refinery-driven petroleum-product exporter. This distinction is important because the same oil-price movement can represent a valuation shock for Norway but a more complex margin, absorption, and product-mix shock for Thailand.

The empirical contribution is a reproducible workspace and rolling end-of-month nowcast experiment for Norway and Thailand. The study assembles official public data, LSEG Workspace market and news data, EIA weekly petroleum indicators, UN Comtrade destination information, freight signals, volatility features, and derived structural-balance features. VP-OilEx-STFN is compared with seasonal naive, SARIMA, SARIMAX, VAR/VECM, MIDAS/ARDL-MIDAS, mixed-frequency dynamic-factor, and baseline structural-fusion benchmarks under the same rolling information set.

The main result is that VP-OilEx-STFN achieves the best accuracy in both country cases. For Norway, the model reduces RMSE by about 15.6 percent relative to the strongest benchmark, the baseline structural-fusion model. This supports the value of separating crude-price, balance, and market-stress channels rather than forcing them into a single linear predictor set. For Thailand, the model reduces RMSE by about 8.1 percent relative to the strongest VAR/VECM benchmark. The smaller but still meaningful gain reflects the more volatile downstream mechanism, where refinery margins and product-export dynamics benefit from structural signals but also require shrinkage toward a stable multivariate anchor.

The results lead to three broader conclusions. First, mixed-frequency information has economic value for oil-export monitoring when it is handled with real-time discipline. Second, structural design matters: country gates and petroleum-specific feature blocks are more credible than a generic black-box sequence model. Third, strong classical benchmarks remain essential. The proposed method is useful not because it ignores econometric structure, but because it learns flexible corrections around transparent benchmark anchors and economically meaningful channels.

The study also has limitations. The current empirical exercise is an end-of-month nowcast rather than a full intramonth vintage experiment. Daily and weekly signals are transformed into month-end information blocks, so the paper does not yet quantify how forecast accuracy improves from day 5 to day 10, day 15, day 20, and month end. Some data blocks, especially freight, news, and destination-market information, are architecture-ready but have shorter or less uniform histories than the official monthly target series. These limitations are handled conservatively in the benchmark design, but they remain important for future deployment.

The next step is therefore clear: extend the rolling evaluation to multiple within-month vintages and report block-level ablations. That extension would show when each information source becomes useful, whether weekly releases materially update the forecast, and how much the country gate shifts as new data arrive. More broadly, the framework can be transferred to other delayed strategic trade flows, including LNG, refined products, petrochemicals, electricity flows, agricultural commodities,

and critical minerals. The main lesson is that AI-based economic monitoring should be structured, auditable, and benchmarked against strong classical alternatives, not treated as a black-box replacement for economic reasoning.

Acknowledgements. This work was supported by Rajamangala University of Technology Suvarnabhumi (RMUTSB) under the Fundamental Fund (FF) scheme, Fiscal Year 2026, through the research project entitled "Application of a Hybrid Model Using Decomposition Method Combined with ARIMAX and Neural Networks for Forecasting Thailand's Oil Exports." The authors also acknowledge the public statistical agencies and data providers cited in the manuscript. The authors are responsible for all analysis, interpretation, and conclusions.

Declarations

Funding. This work was supported by Rajamangala University of Technology Suvarnabhumi (RMUTSB) under the Fundamental Fund (FF) scheme, Fiscal Year 2026, through the research project entitled "Application of a Hybrid Model Using Decomposition Method Combined with ARIMAX and Neural Networks for Forecasting Thailand's Oil Exports."

Competing interests. The authors declare that they have no competing interests.

Ethics approval and consent to participate. Not applicable. This study uses aggregate public and licensed economic, market, trade, and news data and does not involve human participants, human data, human tissue, animals, or clinical data.

Consent for publication. Not applicable.

Data availability. The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Materials availability. Not applicable.

Code availability. The project workspace contains scripts for data construction, feature engineering, benchmark estimation, proposed-model evaluation, figure generation, and table generation. The code is available from the corresponding author on reasonable request, subject to institutional and third-party data-license restrictions.

Author contribution. U.I., N.B., and P.M. contributed to the study conception, research design, interpretation of results, manuscript review, and manuscript editing. P.M. led the methodological design, model specification, empirical workflow, and manuscript preparation. U.I. and N.B. contributed to data organization, validation, literature positioning, and result interpretation. All authors read and approved the final manuscript.

References

[1] Giannone, D., Reichlin, L., Small, D.: Nowcasting: The real-time informational content of macroeconomic data. Journal of Monetary Economics **55**(4), 665–676 (2008)

[2] Ghysels, E., Santa-Clara, P., Valkanov, R.: The midas touch: Mixed data sampling regression models. Technical Report 2004s-20, CIRANO (2004)

[3] Marcellino, M., Schumacher, C.: Factor midas for nowcasting and forecasting with ragged-edge data: A model comparison for german gdp. Oxford Bulletin of Economics and Statistics **72**(4), 518–550 (2010)

[4] Schorfheide, F., Song, D.: Real-time forecasting with a mixed-frequency var. Journal of Business and Economic Statistics **33**(3), 366–380 (2015)

[5] Lim, B., Arik, S.O., Loeff, N., Pfister, T.: Temporal fusion transformers for interpretable multi-horizon time series forecasting. arXiv preprint arXiv:1912.09363 (2019)

[6] Deforce, B., Baesens, B., Diels, J., Asensio, E.S.: Multimix tft: A multi-task mixed-frequency framework with temporal fusion transformers. In: Proceedings of The 2nd Conference on Lifelong Learning Agents. Proceedings of Machine Learning Research, vol. 232, pp. 586–600 (2023)

[7] Challu, C., Olivares, K.G., Oreshkin, B.N., Ramirez, F.G., Canseco, M.M., Dubrawski, A.: N-hits: Neural hierarchical interpolation for time series forecasting. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 37, pp. 6989–6997 (2023)

[8] Nie, Y., Nguyen, N.H., Sinthong, P., Kalagnanam, J.: A time series is worth 64 words: Long-term forecasting with transformers. In: International Conference on Learning Representations (2023)

[9] Wu, H., Hu, T., Liu, Y., Zhou, H., Wang, J., Long, M.: Timesnet: Temporal 2d-variation modeling for general time series analysis. In: International Conference on Learning Representations (2023)

[10] Liu, Y., Hu, T., Zhang, H., Wu, H., Wang, S., Ma, L., Long, M.: itransformer: Inverted transformers are effective for time series forecasting. In: International Conference on Learning Representations (2024)

[11] Jin, M., Wang, S., Ma, L., Chu, Z., Zhang, J.Y., Shi, X., Chen, P.-Y., Liang, Y., Li, Y.-F., Pan, S., Wen, Q.: Time-llm: Time series forecasting by reprogramming large language models. In: International Conference on Learning Representations (2024)

[12] Zeng, A., Chen, M., Zhang, L., Xu, Q.: Are transformers effective for time series forecasting? In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 37, pp. 11121–11128 (2023)

[13] Arslanalp, S., Choi, S.M., Kamali, P., Koepke, R., McKetty, M., Ruta, M., Saraiva, M., Sozzi, A., Verschuur, J.: Nowcasting global trade from space.

Technical Report WP/25/93, International Monetary Fund (2025)

[14] Arslanalp, S., Exton, O., Gao, C., Kamali, P., Saraiva, M., Sozzi, A., Verschuur, J.: Nowcasting country-level trade estimates using imf portwatch. Technical Report WP/2026/099, International Monetary Fund (2026)

[15] Statistics Norway: External trade in goods. Updated May 15, 2026 (2026). https://www.ssb.no/en/utenriksokonomi/utenrikshandel/statistikk/utenrikshandel-med-varer

[16] Statistics Norway: 12780: Exports of crude oil before and after processing in Great Britain, by commodity group, country, contents and year (2026). https://www.ssb.no/en/statbank/table/12780

[17] Norwegian Petroleum: Historical production. Updated March 3, 2026 (2026). https://www.norskpetroleum.no/en/facts/historical-production/

[18] Energy Policy and Planning Office: Petroleum Statistic (2026). https://www.eppo.go.th/index.php/en/en-energystatistics/petroleum-statistic

[19] Energy Policy and Planning Office: Petroleum Price Statistic (2026). https://www.eppo.go.th/index.php/en/en-energystatistics/petroleumprice-statistic

[20] U.S. Energy Information Administration: Petroleum and Other Liquids Data and Weekly Petroleum Status Report (2026). https://www.eia.gov/petroleum/data.php

[21] LSEG: Workspace for Energy Commodities and Historical Market Data (2026). https://www.lseg.com/en/data-analytics/products/workspace/commodities/energy

[22] LSEG Developer Community: LSEG Data Library for Python: News Headlines (2026). https://developers.lseg.com/en/article-catalog/article/lseg-data-library-for-python--news-pagination

[23] Baltic Exchange: Tanker Market Information and Clean and Dirty Tanker Routes (2026). https://www.balticexchange.com/en/data-services/market-information0/tankers-services.html

[24] United Nations Comtrade: UN Comtrade Trade Data and Public API (2026). https://comtrade.un.org/

[25] Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y.: Graph attention networks. arXiv preprint arXiv:1710.10903 (2018)

[26] Diebold, F.X., Mariano, R.S.: Comparing predictive accuracy. Journal of Business and Economic Statistics **13**(3), 253–263 (1995)