

Forecasting Airline Stock Volatility with Oil Futures

Volatility and Oil News Sentiment

Group Number: 23

Student Name: Blake Stanley

Student UT EID: bks2356

Student Name: Shivsagar Palla

Student UT EID: sp56633

Student Name: Raghuvendra Chowdhry

Student UT EID: rbc993

Date: February 27, 2026

Goal (Thesis)

The goal of this project is to investigate whether volatility in crude oil futures, as measured by the CBOE Crude Oil Volatility Index (OVX), and oil market news sentiment, captured by the Text Oil Sentiment Indicator (TOSI), can predict stock price volatility for major U.S. airlines (AAL, DAL, UAL, LUV) and the JETS ETF. We hypothesize that increases in OVX and negative shifts in TOSI lead to increased airline stock volatility, with the strongest predictive power at a 1-month lag. Combining OVX with TOSI should yield a more accurate model than either predictor alone.

Description of the Datasets

Size of the Datasets

The project uses seven datasets: daily Bloomberg implied volatility (IV) estimates for five airline equities, the daily CBOE Crude Oil Volatility Index (OVX), and the monthly Text Oil Sentiment Indicator (TOSI).

Dataset	Rows	Columns	Date Start	Date End
AAL IV	3,089	1	2013-12-09	2026-02-27
DAL IV	4,738	1	2007-05-01	2026-02-27
UAL IV	3,875	1	2010-10-01	2026-02-27
LUV IV	11,511	1	1980-07-28	2026-02-27
JETS IV	2,724	1	2015-04-30	2026-02-27
OVX	4,905	1	2007-05-10	2026-02-25
TOSI	526	1	1982-01-01	2025-10-01

Column Names

Each airline dataset has a Date index and a single implied volatility column (e.g., AAL_Volatility, DAL_Volatility, UAL_Volatility, LUV_Volatility, JETS_Volatility). The OVX

dataset has a Date index and OVX column. The TOSI dataset has a Date index and TOSI column.

Classification of Each Column

All Date indices are ordinal (ordered temporal). All value columns (AAL_Volatility, DAL_Volatility, UAL_Volatility, LUV_Volatility, JETS_Volatility, OVX, TOSI) are continuous quantitative variables. There are no categorical or discrete variables in any dataset.

Discrete and Categorical Variables List and Distribution

Not applicable. All datasets contain only continuous quantitative variables with no categorical data. There are no discrete or categorical variables.

Distribution of Quantitative Variables

The table below summarizes the descriptive statistics for each continuous variable. Airline volatilities are expressed as implied volatility percentages (e.g., 39.55 represents 39.55% annualized implied volatility). OVX is measured in index points representing implied volatility percentage. TOSI is a normalized sentiment index with a mean of approximately zero.

Variable	Min	Max	Range	Median	Mean	Std Dev	Units
AAL Vol	12.41	120.08	107.67	36.34	39.55	13.09	Ann. vol
DAL Vol	19.13	278.83	259.70	39.49	47.71	26.97	Ann. vol
UAL Vol	20.32	310.13	289.81	41.38	44.42	17.74	Ann. vol
LUV Vol	9.92	205.76	195.85	35.15	37.09	12.10	Ann. vol
JETS Vol	10.95	207.44	196.50	28.48	32.66	15.44	Ann. vol
OVX	14.50	325.15	310.65	35.14	38.41	16.94	Index pts
TOSI	-4.0751	2.0867	6.1618	0.1220	~0.00	0.9458	Index

Data Wrangling

List Missing Data

The airline IV datasets contain some missing values: AAL has 13 missing (0.42%), DAL has 6 missing (0.13%), UAL has 1 missing (0.03%), LUV has 3,533 missing (30.69%), and JETS has 17 missing (0.62%). These correspond to days without enough call option volume for implied volatility to be calculated, or early periods after listing with insufficient options data. TOSI has zero missing values. The OVX dataset contains 174 missing values, corresponding to U.S. market holidays when the CBOE was closed and no index was published.

Replacement of Missing Data

For the airline IV datasets (AAL, DAL, UAL, JETS), missing values were backfilled using the next available observation, since the data is ordered from recent to old. Any remaining NaN rows (at the start of each dataset where no prior data exists for backfill) were dropped. For LUV, the same process was applied, but a larger number of rows were lost due to a long stretch of no options data in the early years (the dataset starts in 1980, before widespread option availability). For OVX, the 174 missing values correspond to non-trading holidays and were simply removed. No imputation was applied to TOSI, which had no missing values.

Deletion of Rows and Reasons

The 174 OVX rows with missing values were removed (market holidays). For airline IV datasets, after backfilling, remaining NaN rows at the start of each series were dropped: AAL lost 13 rows, DAL lost 6, UAL lost 1, JETS lost 17, and LUV lost 3,533 rows (due to insufficient early options volume). These are structural absences, not random missingness.

For the airline IV datasets, missing values were backfilled using the next available observation because the data is ordered recent-to-old, so backfill uses past data to fill isolated gaps without leaking future information. The remaining NaN rows at the start of each series (where no prior

data exists) were dropped since forward-filling would introduce look-ahead bias into a prediction model. The 174 missing OVX rows were removed because they correspond to scheduled market holidays when the CBOE did not publish a value, and since the airline datasets also have no observations on these dates, dropping them creates no alignment issues.

Check That Data Is in the Correct Range

All airline volatility values are non-negative, as required for implied volatility estimates. All OVX values are positive (minimum 14.50), consistent with the CBOE volatility index. TOSI values range from -4.08 to 2.09, which is expected for a normalized sentiment index. No out-of-range values were detected in any dataset.

Check That Data Is in the Correct Format

All Date indices are parsed as DatetimeIndex objects. All value columns (volatilities, OVX, TOSI) are stored as float64. Data types are correct and no format conversions were necessary.

Check for Duplicates

No duplicate dates were found in any of the seven datasets. No rows were removed for duplication.

Additionally, since TOSI is a monthly indicator while OVX and airline volatilities are daily, we created monthly aggregated DataFrames by computing the monthly mean of each daily series. This enables correlation analysis and model development on a common time axis. The fully overlapping monthly dataset (all seven columns non-null) spans May 2015 to October 2025 with 126 observations. Pairwise analyses use the maximum available overlap for each pair.

Preliminary Statistical Analysis

Pearson Correlation Coefficients

We computed Pearson correlation coefficients among all quantitative variables on the monthly aggregated data. Pairwise correlations are calculated using the maximum available data for each pair to avoid unnecessarily discarding data.

Key predictor-target correlations (OVX vs. airline volatilities):

OVX shows positive correlation with all airline volatility measures, confirming that oil market uncertainty transmits to airline stocks. The strongest relationship is OVX vs. JETS ($r = 0.81$, $N = 130$), followed by OVX vs. LUV ($r = 0.80$, $N = 226$), OVX vs. UAL ($r = 0.72$, $N = 185$), OVX vs. DAL ($r = 0.57$, $N = 226$), and OVX vs. AAL ($r = 0.52$, $N = 147$).

Key predictor-target correlations (TOSI vs. airline volatilities):

TOSI is negatively correlated with airline volatilities, as expected: positive oil sentiment is associated with lower airline stock volatility. The strongest is TOSI vs. JETS ($r = -0.46$, $N = 126$), followed by TOSI vs. UAL ($r = -0.42$, $N = 181$), TOSI vs. AAL ($r = -0.39$, $N = 143$), TOSI vs. LUV ($r = -0.37$, $N = 380$), and TOSI vs. DAL ($r = -0.27$, $N = 222$).

OVX vs. TOSI:

OVX and TOSI have a moderately strong negative correlation ($r = -0.61$, $N = 222$), confirming that higher oil volatility is associated with more negative oil market sentiment.

Cross-airline correlations:

The airline volatility measures (excluding AAL) are very highly correlated with each other ($r = 0.79$ to 0.97), reflecting strong common sector exposure to macroeconomic and energy market conditions. AAL shows notably weaker correlations with the other airlines ($r = 0.23$ to 0.36), possibly due to its shorter data history and different implied volatility dynamics.

Best Fit Among Highly Correlated Quantities

For all pairs with $|r| > 0.6$, we tested linear, quadratic, and cubic polynomial fits. The cubic fit achieved the highest R-squared in every case, though the improvement over linear was modest for most pairs. This suggests mild non-linearity in the relationships, which motivates the use of non-linear machine learning models (XGBoost/Random Forest) in addition to linear MIDAS regression.

Selected best-fit results:

Variable Pair	R ² Linear	R ² Quad	R ² Cubic	Best Fit
DAL vs JETS	0.9048	0.9096	0.9116	Cubic
DAL vs UAL	0.9345	0.9447	0.9447	Cubic
OVX vs JETS	0.6534	0.6726	0.6984	Cubic
OVX vs LUV	0.6421	0.6552	0.6868	Cubic
OVX vs TOSI	0.3713	0.3801	0.3879	Cubic

Machine Learning Models

Model 1: MIDAS Regression (Mixed Data Sampling)

MIDAS (Mixed Data Sampling) regression is designed for settings where predictors and the dependent variable are sampled at different frequencies. In our case, OVX is daily, TOSI is monthly, and airline volatility is daily. Rather than averaging all variables to the lowest common frequency (which loses information), MIDAS uses parameterized weighting functions (Beta or Almon lag polynomials) to optimally aggregate high-frequency predictors.

We will estimate MIDAS regressions with the following specifications: (1) OVX only as predictor, (2) TOSI only, and (3) both OVX and TOSI combined. We will test multiple lag horizons (1-month, 3-month, 6-month) to identify the most informative lead-lag structure. Model performance will be evaluated via out-of-sample RMSE using a rolling-window forecast scheme and directional accuracy (percentage of correctly predicted volatility increases/decreases).

Model 2: XGBoost / Random Forest

XGBoost and Random Forest are tree-based ensemble methods that can capture non-linear relationships and feature interactions without requiring explicit specification. Given that our best-fit analysis found cubic relationships outperforming linear ones, these models are well-suited to exploit the non-linearity in the OVX-airline volatility relationship.

Features will include: lagged monthly OVX statistics (mean, max, min, standard deviation of daily values within the month), lagged TOSI values, and lagged airline volatility (autoregressive component). The target variable is next-month average airline volatility. We will use an 80/20 temporal train-test split (no shuffling, respecting time-series ordering) with time-series cross-validation for hyperparameter tuning. Feature importance analysis from both models will reveal which predictors are most influential and at what lags.

Narration

Following the storytelling principles outlined by Kosara and Mackinlay (and inspired by Hans Rosling's dynamic data presentations), our 5-minute narration proceeds as follows:

Opening — Setting the Stage (0:00–0:45):

Airlines spend 25–35% of operating costs on fuel, making them uniquely vulnerable to oil market disruptions. When oil volatility spikes, airline margins compress, hedging costs rise, and investor uncertainty increases. The question: can we predict airline stock turbulence by monitoring oil markets?

Introducing the Data (0:45–1:30):

We present three data streams: (1) daily implied volatility (IV) for AAL, DAL, UAL, LUV, and JETS — capturing the intensity of daily price swings; (2) OVX, the CBOE's oil volatility fear gauge, published daily since May 2007; and (3) TOSI, a novel monthly sentiment indicator built from NLP analysis of 6 million Financial Times, Reuters, and Independent articles mentioning oil prices (1982–2025).

The Oil–Airline Connection (1:30–2:30):

Our correlation analysis reveals OVX positively correlates with all airline volatilities ($r = 0.52$ to 0.81), confirming oil uncertainty transmits to airline stocks. TOSI is negatively correlated ($r = -0.27$ to -0.46): positive sentiment calms airline markets. The airline stocks themselves are very highly correlated ($r > 0.76$), moving as a sector.

The Mixed-Frequency Challenge (2:30–3:15):

OVX moves daily, TOSI moves monthly. Naive monthly averaging loses the day-to-day OVX dynamics that could be most predictive. This motivates MIDAS regression, which preserves high-frequency information via optimized weighting schemes. We complement this with XGBoost to capture non-linear interactions the cubic-fit analysis suggests are present.

The Prediction Question — Climax (3:15–4:15):

Can combining traditional volatility data (OVX) with NLP-derived sentiment (TOSI) beat models using either alone? At what lag horizon is prediction strongest? Gifuni's research shows TOSI outperforms other sentiment proxies especially during crises (COVID-19, Russia-Ukraine). We test whether this advantage extends to airline volatility forecasting.

Expected Resolution (4:15–5:00):

We hypothesize the combined OVX + TOSI model outperforms, particularly during crisis periods when sentiment data captures market shifts faster than traditional volatility indices. The practical implication: portfolio managers could use this combined signal to adjust airline exposure ahead of volatility moves, and algorithmic trading strategies could exploit the cross-asset predictive channel.