

07_business_interpretation_and_forecast

January 29, 2026

1 Purpose

Summarise forecasting results in business terms and demonstrate how return and volatility forecasts inform practical risk and decision-making.

2 Imports

```
[1]: import pandas as pd
import numpy as np
from IPython.display import Image, display

[2]: # Configuration for clean output
pd.set_option('display.max_columns', None)
pd.set_option('display.float_format', '{:.6f}'.format)
```

3 Returns

3.1 Accuracy comparison

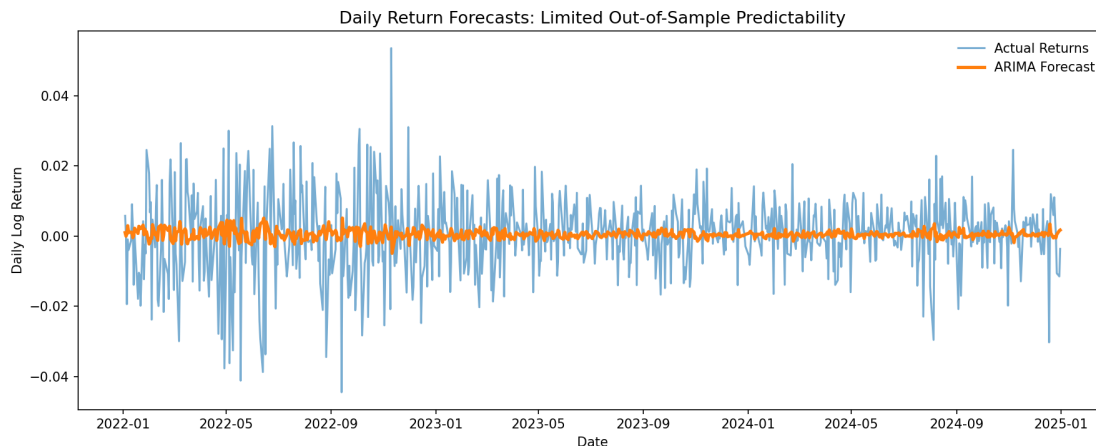
```
[7]: return_results = pd.read_csv("../outputs/tables/return_model_comparison.csv")
return_results.style.hide(axis='index')
```

```
[7]: <pandas.io.formats.style.Styler at 0x1b56ea95550>
```

No material improvement over baseline at a daily horizon, consistent with return noise dominating signal.

3.2 Visual comparison

```
[4]: display(Image(filename='../outputs/figures/return_forecast_comparison.png'))
```



Forecasts collapse toward zero, highlighting the dominance of noise over signal.

Out-of-sample predictability of daily returns is negligible.

3.3 Interpretation

The ARIMA model provides no material improvement over the Baseline Mean. The MAE/RMSE metrics are virtually identical.

This confirms that daily price history contains negligible predictive power for future direction.

Forecasts should not be relied upon for directional trading (i.e., Long or Short). Transaction costs would almost certainly erode any theoretical edge seen in the 4th decimal place.

4 Volatility

4.1 Accuracy comparison

```
[11]: vol_results = pd.read_csv("../outputs/tables/volatility_model_comparison.csv")
      vol_results.style.hide(axis='index')
```

```
[11]: <pandas.io.formats.style.Styler at 0x1b56e481590>
```

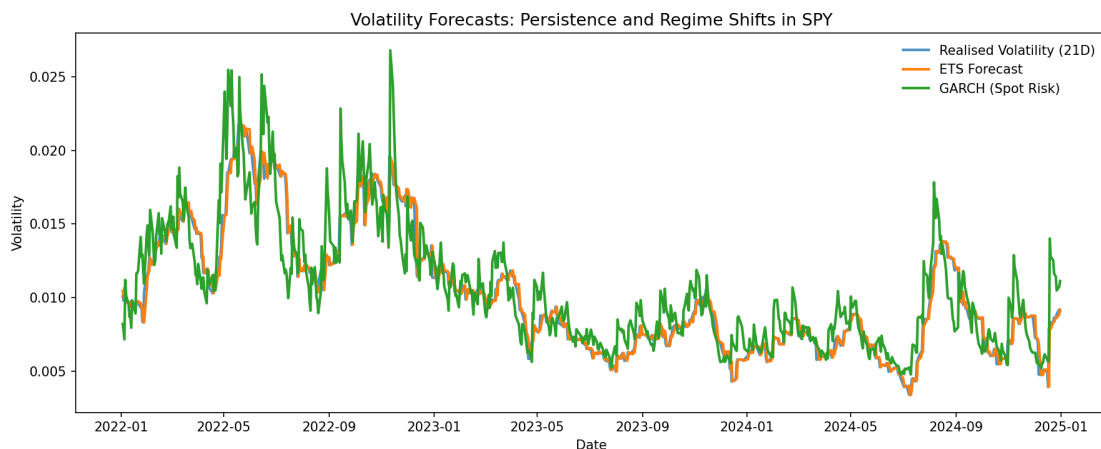
Naive and ETS models yield nearly identical errors because, at a one-day horizon, both effectively follows a random walk: $\hat{\sigma}_t = \sigma_{t-1}$, benefiting from the extreme autocorrelation in the 21-day realised volatility series.

Raw GARCH forecasts exhibit the highest error metrics because they estimate daily spot risk, whereas the target is a smoothed 21-day average. This frequency mismatch statistically penalises GARCH for its inherent reactivity.

Smoothing GARCH forecasts to a 21-day window reduces the annualised MAE from 2.29% to 1.55%, confirming that GARCH captures the structural level of risk effectively once the noise of different measurement horizons is mitigated.

4.2 Visual comparison

```
[6]: display(Image(filename='../outputs/figures/volatility_forecast_comparison.png'))
```



Volatility is persistent, not random.

Both ETS and GARCH capture the level of volatility far better than any return model captures direction.

However,

ETS behaves like a smoother:

- Excellent at tracking medium-term regimes
- Low noise, high stability
- Lags during sudden jumps, but doesn't overreact

GARCH behaves like a shock detector:

- Responds immediately to large return shocks.
- Captures clustering and conditional heteroskedasticity.
- Noisier and less accurate in RMSE terms because it forecasts risk, not smooth averages.

4.3 Interpretation

Volatility is forecastable in structure, even if not precisely timed.

These forecasts are suitable for risk limits, position sizing, stress testing, and capital allocation.

GARCH captures conditional variance dynamics not visible in smoothing models, which is valuable for modelling conditional variance dynamics relevant to risk management.

Volatility is highly forecastable in magnitude (level), even if the specific timing of shocks is stochastic.

4.3.1 The Accuracy Winner

The Naive and ETS models perform best on RMSE because they track the high persistence of the target variable (21-day Realised Volatility).

They are better at Reporting & Accounting. If we need to predict “What will the risk report say tomorrow?”, use ETS.

4.3.2 The Utility Winner

GARCH has higher RMSE because it reacts to structural shocks rather than just averaging the past.

It is the best for Risk Management & Trading. If we want to answer “Is the market crashing now?”, use GARCH.

5 Conclusion

Daily return forecasting is futile with simple time-series models. The signal-to-noise ratio is too low.

Volatility forecasting is highly effective. We can reliably predict the risk regime (High/Low volatility) even if we cannot predict the next day’s price.

Model Selection:

- Use ETS for smooth, stable estimates of “average” volatility (e.g., setting longer-term capital buffers).
- Use GARCH for reactive, tactical estimates of “spot” volatility (e.g., daily VaR limits or adjusting leverage).

6 Limitations

Analysis considers only SPY history. It ignores macro factors (VIX, Interest Rates) that drives volatility.

The GARCH model assumes parameters (α, β) are constant over 2022–2024. In reality, market structure shifts (regime changes) often require re-estimating parameters.

Analysis is limited to one-step-ahead (Daily). Long-term forecasts (Monthly/Quarterly) might favor different models.