SP500 Time-Series Forecasting

A Box-Jenkins and GARCH Approach

Justin Wang and Remi Liu

STAT 443
2025

November 2025

Outline

Problem Statement

- **Objective**: Forecast SP500 returns and volatility using time-series methods
- **Challenge**: Financial time series exhibit:
  - Non-stationarity
  - Volatility clustering
  - Complex dependencies
- **Approach**:
  - Box-Jenkins methodology (AR, MA, ARIMA) for returns
  - GARCH/ARCH models for volatility

Why It Matters

Investment Management

- Portfolio allocation
- Market timing
- Risk-return tradeoffs

Risk Management

- Value-at-Risk (VaR)
- Position sizing
- Hedging strategies

Derivatives Pricing

- Option pricing requires volatility forecasts
- Black-Scholes and other models

Data Sources

- **Period**: October 1, 2015 to October 30, 2025
- **Observations**: 2,631 daily observations
- **Primary Data**:
    - SP500 ETF (SPY): Daily closing prices
    - VIX Index: Market volatility expectations
    - 10-Year Treasury Yield: Risk-free rate proxy
    - High-Yield Credit Spread: Credit risk indicator
- **Target Variable**: Daily log returns

$$r_t = \log(P_t) - \log(P_{t-1})$$

Data Characteristics

Descriptive Statistics

- Mean return: 0.04% daily
- Std dev: 1% daily
- Distribution: Fat-tailed
- Volatility clustering: Present

Stationarity Tests

- ADF test: $p < 0.05$
- KPSS test: $p > 0.05$
- **Result**: Returns are stationary (d=0)
- No differencing required

Feature Engineering

- Created **36 potential predictors** from raw market data
- Categories:
  - Lagged returns (R_lag1, R_lag2, R_lag5)
  - Realized volatility (5-day, 20-day)
  - Technical indicators (RSI, moving averages)
  - Cross-asset features (VIX ratios, yield curves)
  - Interaction terms
- **Variable Selection**: Elastic Net regularization
  - Selected **8 predictors** (22% selection rate)
  - Stability threshold: ≥ 0.6

Box-Jenkins Methodology

Three-Step Procedure:

1. **Identification**
    - ACF/PACF analysis
    - Stationarity testing
2. **Estimation**
    - Grid search: AR(p), MA(q), ARIMA($p, 0, q$)
    - Selection criteria: AIC, BIC
3. **Diagnostic Checking**
    - Ljung-Box test (residual autocorrelation)
    - Jarque-Bera test (normality)
    - Residual plots

ARCH(q) Model

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$$

- Models volatility using past squared errors
- Simpler, fewer parameters

GARCH(p,q) Model

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

- Includes lagged variance terms
- Captures volatility persistence
- GARCH(1,1) is standard in finance

Backtesting Procedure

- **Training Set**: 80% (2015-10-01 to 2023-10-23)
- **Test Set**: 20% (2023-10-24 to 2025-10-30)
- **Method**: Rolling-origin expanding window
- **Forecast Horizon**: 1-step-ahead
- **Test Folds**: 528 (one per test observation)
- **No Future Leakage**: Each forecast uses only past data

Evaluation Metrics:

- RMSE, MAE, MAPE
- Directional accuracy
- Diebold-Mariano test

Model Selection: Return Forecasting

Model	RMSE	MAE	MAPE (%)	Dir. Acc. (%)
ARIMA(2,0,2)	**0.0100**	0.00655	6439	49.6
AR(8)	0.0101	0.00653	6969	51.1
MA(2)	0.0101	0.00640	2838	**53.4**

Selected: ARIMA(2,0,2)

- Lowest RMSE (best overall accuracy)
- AIC: -16,287.1, BIC: -16,251.85
- Combines AR and MA components

Model Selection: Volatility Forecasting

Model	Order	AIC	BIC	Mean Vol. (%)
GARCH	**(1,1)**	**-6.65**	**-6.64**	**0.931**
ARCH	(1)	-6.14	-6.14	1.152

Selected: GARCH(1,1)

- Lower BIC (-6.64 vs -6.14)
- Time-varying volatility (0.855% to 0.991%)
- ARCH produces constant volatility (unrealistic)

Model Diagnostics

ARIMA(2,0,2)

- Ljung-Box: Some residual autocorrelation (acceptable)
- Jarque-Bera: Non-normal residuals (expected)
- Residual ACF: Mostly within bands

GARCH(1,1)

- Residuals: White noise
- Squared residuals: No ARCH effects
- Volatility clustering: Captured

Diebold-Mariano Tests:

- No statistically significant differences between AR, MA, ARIMA
- All models have similar predictive power

Return Forecasts

21-Day Forecast (Oct 31 - Nov 28, 2025):

- **Cumulative Return**: +1.12%
- **Annualized**: ∼13.4%
- **Directional**: 81% positive days (17 up, 4 down)
- **Mean Daily Return**: +0.053%
- **Prediction Intervals**: 95% width averages 4.36%

Interpretation:

- Bullish short-term outlook
- Moderate volatility expected
- Wide prediction intervals reflect uncertainty

Volatility Forecasts

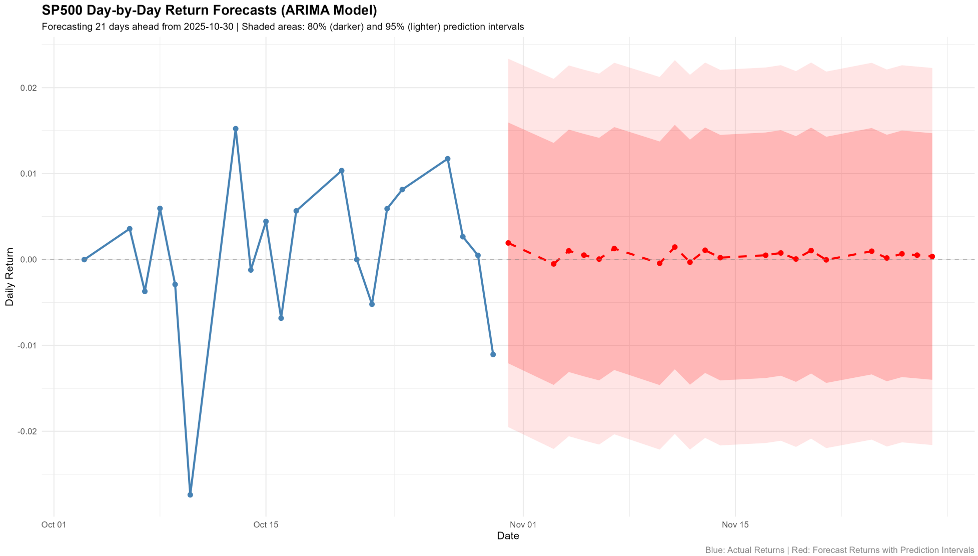
GARCH(1,1) Forecast:

- **Mean Volatility**: 0.931% (vs historical: 0.843%)
- **Range**: 0.855% to 0.991% (time-varying)
- **Change**: +10.4% increase from historical mean
- **Trend**: Gradually increasing over 21 days

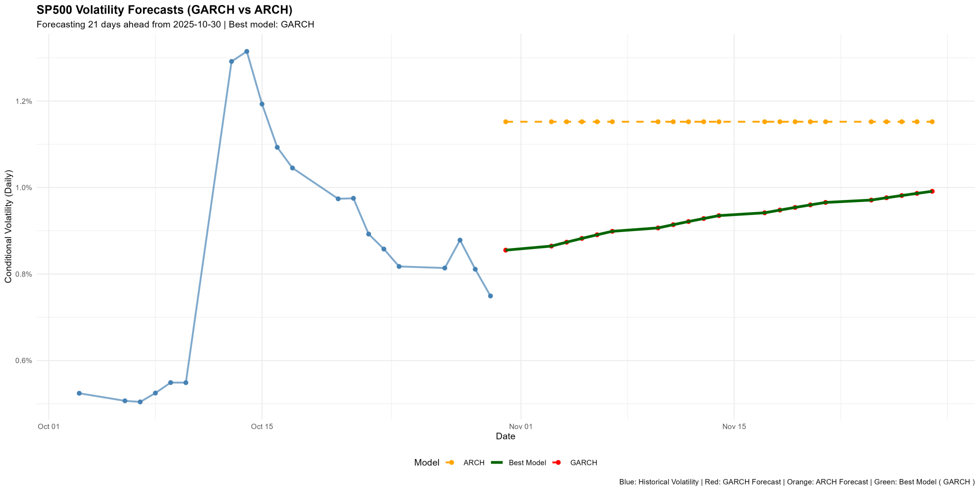
Implications:

- Increasing market risk ahead
- Consider reducing position sizes
- Increase hedging activities
- Adjust VaR calculations upward

Forecast Visualization



Volatility Forecast Visualization



Key Findings

1. **Best Return Model**: ARIMA(2,0,2)
   - RMSE: 0.0100 (best accuracy)
   - All models show similar performance
2. **Best Volatility Model**: GARCH(1,1)
   - BIC: -6.64 (clearly superior to ARCH)
   - Produces realistic time-varying forecasts
3. **Forecast Insights**:
   - Bullish short-term outlook (+1.12% over 21 days)
   - Increasing volatility (+10.4% from historical)

Practical Implications

Investment Strategy

- Maintain equity exposure
- 81% positive days forecast
- But account for uncertainty

Risk Management

- Reduce position sizes
- Increase hedging
- Adjust VaR upward

Caveat: Forecasts are probabilistic, not deterministic. Actual outcomes may differ, especially during unexpected market events.

Limitations and Future Work

Limitations:

- Stationarity assumption may not hold during regime changes
- Linear models may miss non-linear dependencies
- Short-term focus (1-step-ahead forecasts)
- No exogenous variables in pure time-series approach

Future Improvements:

- ARIMAX models with external predictors
- GARCH extensions (GJR-GARCH, EGARCH)
- Regime-switching models
- Machine learning comparisons (LSTM, XGBoost)

Summary

- Successfully applied Box-Jenkins and GARCH methodologies to SP500 forecasting
- Selected ARIMA(2,0,2) for returns and GARCH(1,1) for volatility
- Generated practical forecasts with prediction intervals
- Provided actionable insights for investment and risk management