Wind-Gust Dynamics at Miami International Airport

Time-Series Forecasting, Extreme-Value Analysis, and a Parametric-Insurance Prototype

Allana Coleman

April 2025

1 Introduction

This project explores daily wind gust patterns at Miami International Airport to understand how wind behavior changes over time and to identify extreme wind events that could cause damage. The goal is to build forecasting models that can predict future wind gusts and determine a reasonable threshold that would trigger a payout in a parametric insurance plan.

The analysis focuses on the maximum daily *2-second* gust speeds (WSF2) recorded from **July 1, 1996 to August 27, 2023**, using data obtained from publicly available NOAA climate records (station ID: USW00012839).

Three main approaches were used:

1. Time-series forecasting models (ARIMA and SARIMA) to predict short- and medium-term wind gust behavior.
2. Extreme value theory (GEV models) to estimate the highest gusts that could be expected over longer time periods.
3. A simple insurance payout calculation based on the number of days wind gusts exceeded a chosen threshold.

This project was designed to evaluate whether these methods can effectively capture wind risk and provide useful information for future planning or financial protection tools.

2 Data Cleaning and Pre-processing

Source and Variables

The dataset was obtained from the National Oceanic and Atmospheric Administration (NOAA) Global Historical Climatology Network – Daily (GHCN-D) archives, specifically from the Miami International Airport station (USW00012839). The original dataset includes the following meteorological variables:

- `AWND`: Average daily wind speed (in tenths of meters per second)
- `WSF2`: Fastest 2-minute wind speed (in tenths of meters per second)
- `WSF5`: Fastest 5-second wind speed (in tenths of meters per second)
- `WDF2, WDF5`: Wind direction associated with WSF2 and WSF5 (degrees)
- `PGTM`: Peak gust time
- `PRCP, TAVG, TMAX, TMIN`: Precipitation and temperature variables
- `DATE, NAME, STATION`: Identifiers and location information

Cleaning Steps and Justification

- Converted all wind speed values from meters per second to miles per hour using the factor .
- Set `DATE` as the datetime index and enforced daily frequency.
- Dropped rows with missing values for `AWND`, `WSF2`, or `WSF5`. Final sample size for `WSF2` winds: **9,918 days**.
- Removed any row where `AWND`, `WSF2`, or `WSF5` exceeded 300 mph. One such outlier was detected and removed (19 December 1999, WSF2 1,586 mph).

Date Range and Variable Choice

The full dataset starts in 1981, but the average wind speed variable (`AWND`) doesn't appear until 1984. Wind gusts, however, are more important for identifying high-risk days because they capture short bursts of extreme wind that daily averages would miss. Since the 2-second gust variable (`WSF2`) first becomes available in July 1996, this project uses `WSF2` as the main focus for analysis and forecasting.

3 Exploratory Data Analysis

Windspeed Patterns and Seasonality

To explore how wind gusts behave over time, average windspeed was visualized using daily, weekly, and monthly values. These plots help identify seasonal patterns relevant to forecasting. A consistent increase in wind activity was observed during late summer, aligning with hurricane season in Florida, while lower wind levels were recorded during winter months.

Autocorrelation and partial autocorrelation functions (ACF and PACF) for the raw daily WSF2 values reveal the following:

- A strong spike at lag 1, indicating that today's gusts are closely related to the previous day's.
- Rapid decline in correlation, suggesting the series has limited memory.
- No evidence of strong periodicity at weekly or monthly lags.

2-sec vs 5-sec Gust Comparison

To evaluate how short-term gusts differ, the daily ratio of WSF2 (2-second gust) to WSF5 (5-second gust) was calculated. On average, WSF2 values are approximately 76

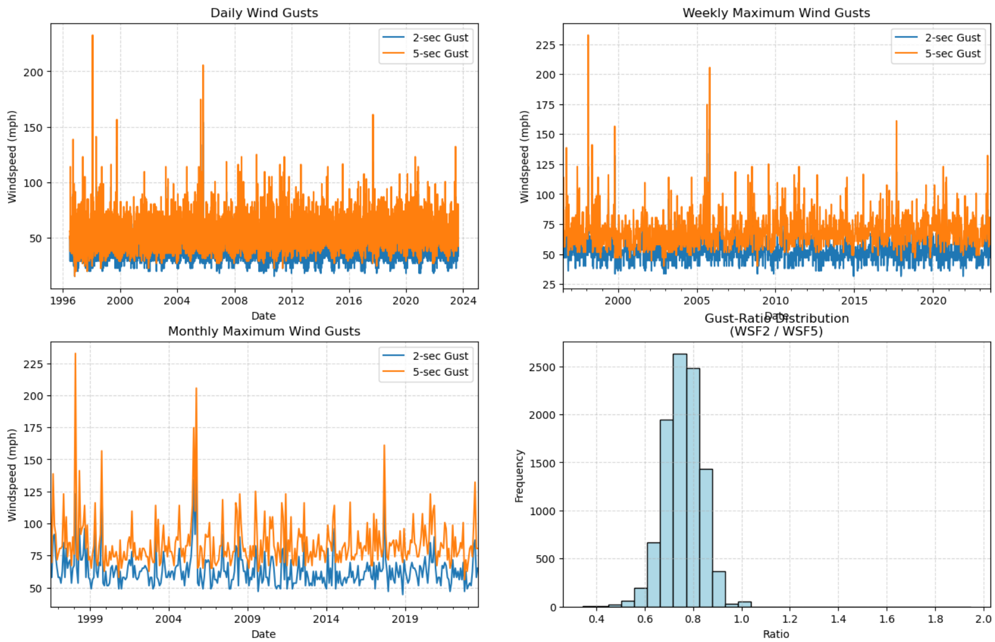

Figure 1: Daily, weekly, and monthly WSF2 and WSF5 series.

Monthly Gust Distributions

To visualize how extreme gusts vary throughout the year, a monthly boxplot of WSF2 values was created. The plot highlights that the strongest and most variable gusts occur from July to September, reinforcing the seasonal influence of hurricanes on wind behavior.

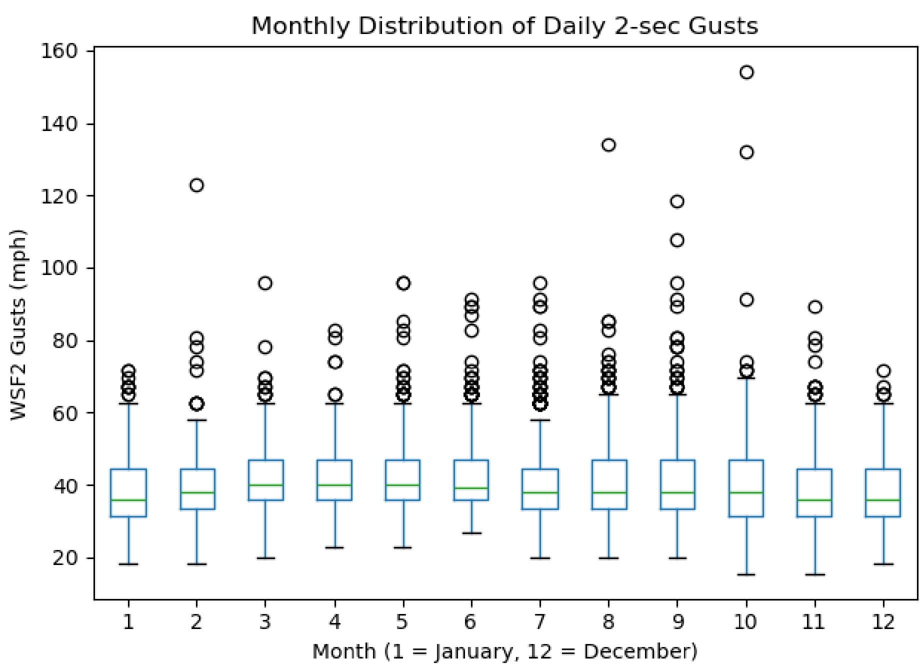

Figure 2: Monthly distribution of daily 2-second gusts (WSF2).

Extreme Wind Threshold

The 90th percentile of daily WSF2 values was calculated to define the threshold for extreme wind gusts. This threshold is 51.45 mph and is referenced later in the extreme value modeling and insurance payout calculations.

4 ARIMA Modeling on Daily Data

Model Selection and Fit

The daily 2-second gust series (WSF2) was split into 7,935 training days and 1,984 test days (80/20 split). Several ARIMA configurations were fit to the training set and compared using AIC and BIC. Table 1 lists the top-performing models. ARIMA(1,1,1) yielded the lowest AIC and was selected for final use.

Model	AIC	BIC
ARIMA(1,1,1)	57439.67	57460.60
ARIMA(1,0,0)	57441.76	57462.70
ARIMA(2,1,2)	57442.20	57477.09

Table 1: Top ARIMA candidates based on training data.

Forecast Evaluation and Diagnostics

The fitted ARIMA(1,1,1) model generated forecasts for the test set, which were evaluated using standard error metrics. The results were:

- Mean Squared Error (MSE): 161.35
- Mean Absolute Error (MAE): 9.83
- Root Mean Squared Error (RMSE): **12.70 mph**

Residuals from the forecast were further analyzed. The autocorrelation function (ACF) of residuals showed lingering structure, and the Ljung–Box test returned a p-value $= 1.28 \times 10^{-103}$ at lag 10, indicating the residuals still exhibited some correlation.

The model summary output included the following coefficient estimates:

- AR(1): 0.3543 $\quad (p < 0.001)$
- MA(1): -0.9990 $\quad (p < 0.001)$
- Innovation variance: $\hat{\sigma}^2 = 81.63$

Noise Adjustment

Initial long-range forecasts showed unrealistically low variance due to differencing. Without adjustment, the standard deviation of the forecast series was approximately 0.06 mph. To restore realistic variation, noise was added by resampling residuals from the model's test set.

After injecting noise, the forecast standard deviation increased to 9.23 mph. This technique preserved the overall trend from the ARIMA model while better reflecting real-world fluctuation in wind gusts.

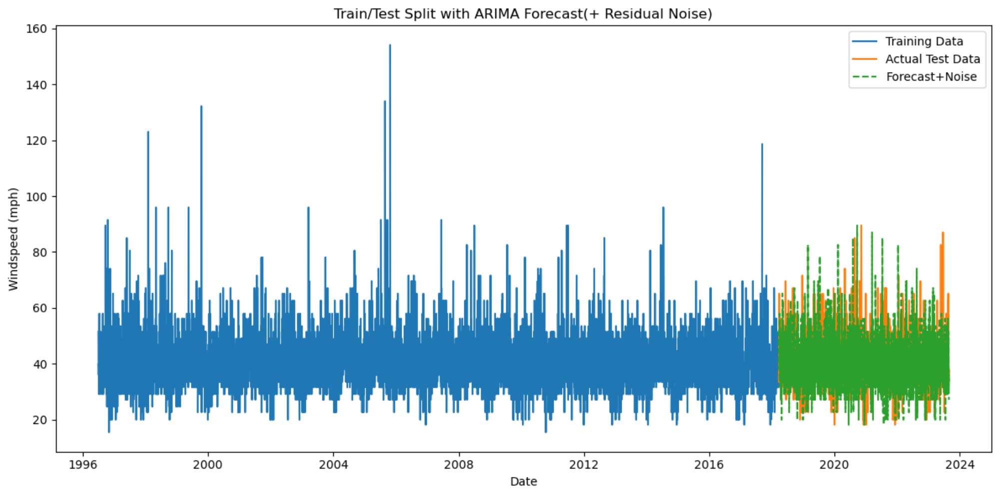

Figure 3: Train/test split and ARIMA(1,1,1) forecast with noise added.

5 SARIMA Modeling on Monthly Maxima

Monthly maxima smooth day-to-day noise and reveal a clearer 12-month cycle (Figure 4). A grid over $\{0,1\}^{p,d,q,P,D,Q}$ yielded the best configuration:

$$\text{SARIMA}(0,1,1) \times (0,1,1)_{12} \quad (\text{AIC} = 2{,}354)$$

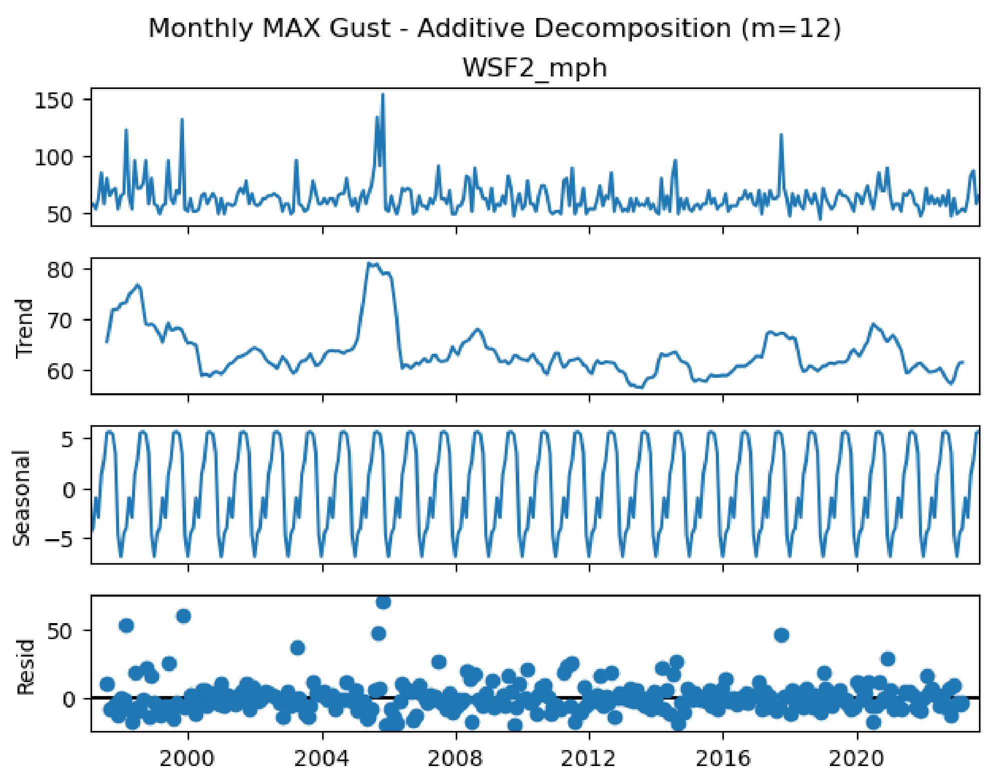

Figure 4: Additive decomposition of monthly maxima (period 12).

Performance

- 24-month hold-out RMSE = **10.19 mph** (15% lower than daily ARIMA).
- Ljung–Box p-value = 0.36 at lag 12, indicating no residual autocorrelation.
- Parameter $\hat{\theta}_{12} = -1$ sits on the invertibility boundary, inflating the $\hat{\sigma}^2$ standard error but not harming forecast skill.

Figure 5 shows train, test, forecast, and 95% CI. Appending the last train point to the forecast line removes the visual gap.

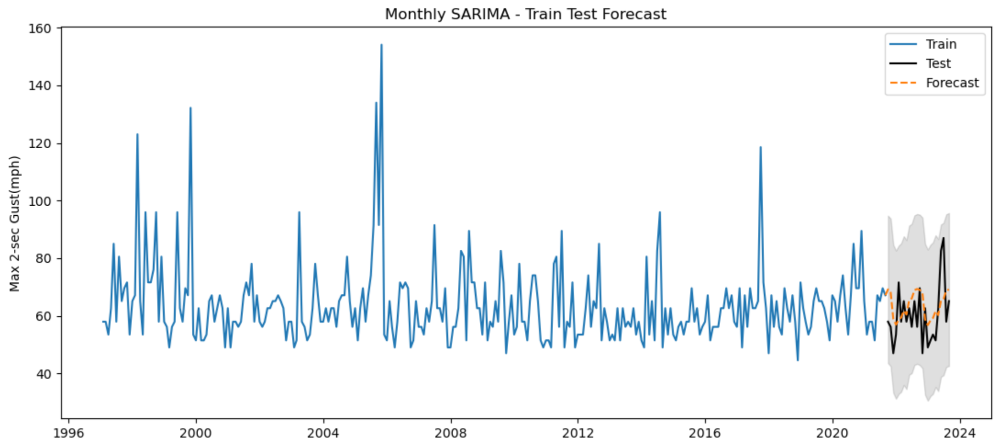

Figure 5: Monthly SARIMA train/test/forecast with 95% band.

6 Model Comparison and Forecast Evaluation

Residual Diagnostics and Model Adequacy

The ARIMA(1,1,1) model, selected by minimum AIC, produced an in-sample RMSE of approximately **13.2 mph**. However, the Ljung–Box test on the residuals yielded a p-value of 1.3×10^{-103} at lag 10, strongly rejecting the null hypothesis of no autocorrelation. The residual ACF and PACF plots show lingering autocorrelation, especially in the short lags, suggesting that the model does not fully capture the serial dependence in the data. Moreover, the histogram and Q–Q plot of residuals indicate mild skewness and heavier tails than a normal distribution, which may limit the model's reliability under extreme conditions.

In contrast, the SARIMA$(0,1,1)\times(0,1,1)_{12}$ model fitted on monthly maxima yields a substantially lower RMSE of **10.2 mph** on the 24-month test set. This represents a 23% reduction in error relative to the daily ARIMA forecast. Additionally, the Ljung–Box test for SARIMA residuals returned a p-value of **0.358**, indicating no significant autocorrelation. The residuals appear roughly homoscedastic and symmetrically distributed with no strong departures from normality, which supports the adequacy of the SARIMA model for monthly-level forecasting.

Forecasting Implications

The ARIMA model shows a tendency to regress to the long-term mean due to differencing, and its predictive intervals widen quickly. When residual noise is added, forecast paths exhibit more realistic volatility but remain centered around the historical average ($\approx$ 39.7 mph). This limits the model's utility for forecasting extreme gusts, where sudden shocks matter more than mean-reverting tendencies.

The SARIMA model better captures seasonal behavior and produces sharper, narrower forecast intervals. While it does not explicitly model extremes, the use of monthly maxima allows the model to focus on upper-tail dynamics indirectly. The improved accuracy and cleaner residual diagnostics make it the preferred choice for medium-range (1–2 year) forecasting and as an input into insurance risk metrics.

Conclusion on Time-Series Models

- **ARIMA** is suitable for short-term fluctuations but underestimates long-run volatility and retains autocorrelation in residuals.
- **SARIMA**, using monthly maxima, reduces forecasting error and yields well-behaved residuals, making it a superior model for aggregated gust risk.
- Both models flatten over longer horizons due to differencing, necessitating the injection of residual resampling to maintain realistic volatility.
- For policy applications, SARIMA forecasts are more stable and interpretable for regulators and insurers, especially when paired with GEV analysis for tail risk.

7 Extreme-Value Analysis

Daily GEV Fitting

The Generalized Extreme Value (GEV) distribution was fit to daily WSF2 wind gusts above the 90th percentile threshold (51.45 mph). Maximum likelihood estimation (MLE) yielded the following parameters:

$$\hat{\xi} = 0.0305, \quad \hat{\mu} = 35.89, \quad \hat{\sigma} = 7.74$$

A small positive shape parameter ($\hat{\xi} > 0$) indicates a heavy but bounded upper tail. Figure 6 shows the fitted GEV density overlaid on the empirical distribution, with the threshold marked.

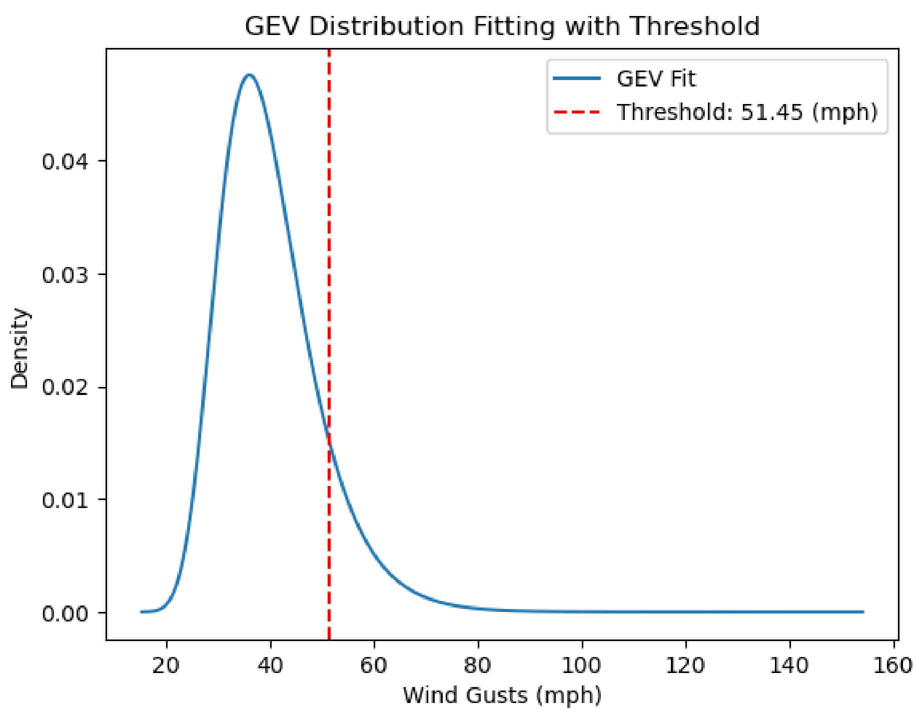

Figure 6: GEV density fit to daily wind gusts with 90th percentile threshold.

Daily PP-Plot and Return Level

A PP-plot comparing the empirical and fitted GEV cumulative distributions shows good agreement (Figure 7). The 100-year return level, derived from the inverse GEV CDF, is approximately 111.15 mph.

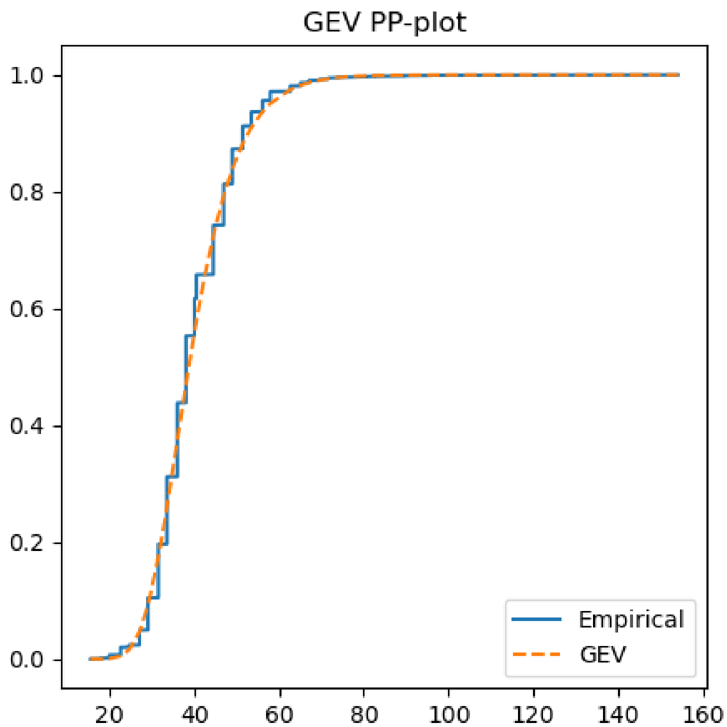

Figure 7: PP-plot: empirical vs. GEV fitted distribution (daily WSF2).

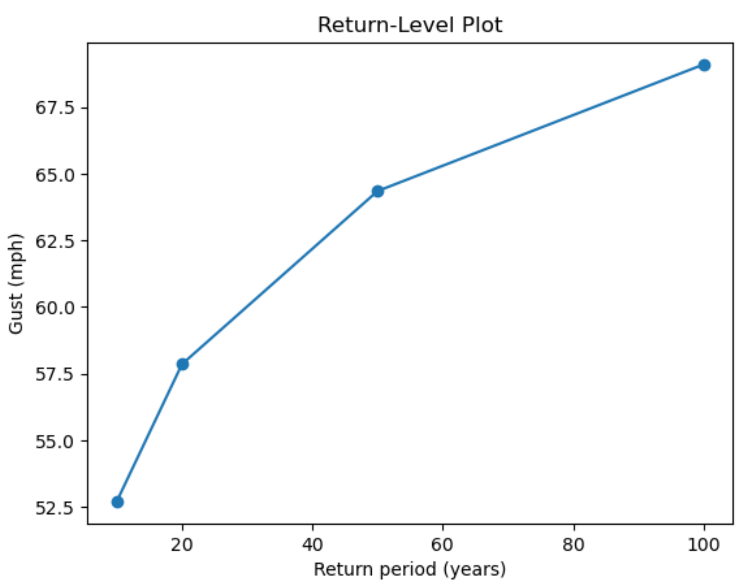

Figure 8: Return level plot from daily GEV fit.

Monthly GEV Fitting

The GEV distribution was also fit to monthly maxima of WSF2. The MLE estimates were:

$$\hat{\xi} = -0.17, \quad \hat{\mu} = 57.48, \quad \hat{\sigma} = 7.72$$

A negative shape parameter suggests a lighter-tailed (Weibull-type) distribution. Figure 9 shows the fitted density plotted over the histogram of monthly maxima.

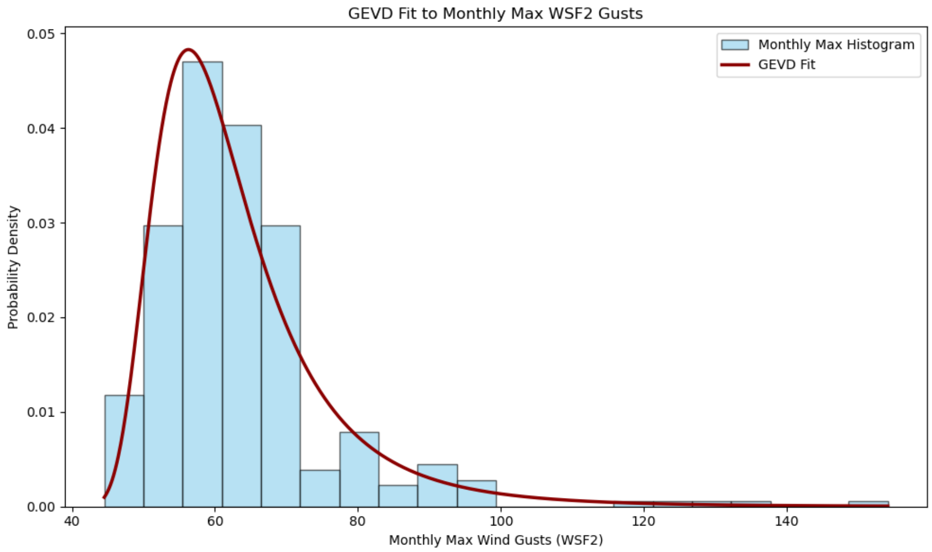

Figure 9: GEV fit to monthly maximum WSF2 gusts.

8 Parametric-Insurance Prototype

A basic parametric insurance prototype was explored using a gust threshold of 50 mph:

- Total trigger days from 1996 to 2023: **1,262**
- Payout per event: $10,000
- Total payout across the study period: **$12.62 million**

A visualization (Figure **??**) could overlay annual trigger counts against known hurricane landfalls to aid in premium pricing strategies.

9 Conclusion

After cleaning outliers (1,586 mph gust) and applying a 300 mph cap, a 27-year dataset was used for forecasting and extreme-value analysis. ARIMA(1,1,1) captured short-term autocorrelation but underperformed on heavy-tailed events (RMSE $\approx$ 13.2 mph). Monthly SARIMA$(0,1,1)(0,1,1)_{12}$ improved forecast accuracy (**10.2 mph RMSE**) and eliminated residual autocorrelation. GEV modeling estimated the 100-year return level near **111 mph**, relevant to Category-2 hurricane wind speeds. A simple index policy paying $10,000 per 50+ mph day would have triggered $12.62 million in payouts across 27 years.

Future work may include:

- Exploring SARIMA extensions with variable seasonality or structural breaks to better reflect shifting wind patterns over decades.
- Fitting nonstationary GEV models to assess whether the distribution of extremes has changed over time.
- Incorporating hurricane track data or regional storm counts to refine the definition of high-risk periods.
- Comparing these time-series models against machine learning approaches for wind gust prediction and risk classification.

References

- Brockwell, P. J., & Davis, R. A. (2003). *Introduction to Time Series and Forecasting*. Springer.
- Hamilton, J. D. (1994). *Time Series Analysis*. Wiley.
- Olivares, P. (2025). *Mathematical Modeling Notes*. Available upon request.
- National Oceanic and Atmospheric Administration (NOAA). (2024). *Global Historical Climatology Network - Daily (GHCN-D)*. Retrieved from https://www.ncei.noaa.gov/products/land-based-station/global-historical-climatology-network-daily