# CAT Bond as an Investment Instrument: a Time Series Analysis

Dennis Goldenberg, Luyang Feng, Christine Lu

## Summary

This project aims to predict catastrophe (CAT) bond market yields using key variables, including expected loss and credit spreads in the non-investment grade bond market. Using a series of autoregressive models, we evaluate the predictive power of these variables, comparing model performance with the Bayesian Information Criterion (BIC) to identify the best-fit model. This analysis provides insights into the factors driving CAT bond yields, with implications for risk management and investment strategies.

To conduct our analysis, we are using data from Artemis (catastrophe bond issuance), FRED (credit spread data), and general market index data. This study will provide insights into the factors influencing cat bond yields and their connection to broader market trends.

## 1. What is a Catastrophe Bond?

A catastrophe (cat) bond is a financial instrument that enables insurers or sponsors to transfer the financial risk of catastrophic events, like natural disasters, to the capital markets. Issued by a Special Purpose Vehicle (SPV), these bonds provide returns to investors through premiums and investment income. If no triggering event occurs, investors receive their principal and interest; if a covered event happens, the bond defaults, and the funds are used to pay insurance claims.

Cat bond pricing is influenced by factors such as reinsurance market conditions, catastrophe modeling, and credit spreads in high-yield bonds. Catastrophe models assess the probability of default for specific risks, aligning premiums with credit spreads for similar bond ratings. While these bonds have high setup costs due to SPV creation and regulatory requirements, multi year agreements help reduce annual expenses. Designed for fat-tail risks, cat bonds provide a unique way to diversify financial portfolios while addressing major disaster risks.

## 2. Exploratory Data Analysis

In our analysis, we aim to model CAT bond yields based on key explanatory variables that could influence their movements.

The primary response variable in our analysis is the Average CAT Bond Market Yield, referred to as "Weekly Average Reinsurance Market Yield (USD)" in the dataset. This variable captures the weighted average annual yield for CAT bonds in the market. It serves as a benchmark for market performance, reflecting investor expectations for return, adjusted for the risk of catastrophic loss. The dataset spans from 2010-10-08 to 2024-10-25. We examined several covariates that potentially influence CAT bond market yields:

1. **Money Market Rate (USD):**
   This variable represents short-term interest rates in the U.S. money market, providing a baseline for risk-free returns. It serves as a component of CAT bond yields, reflecting the underlying economic environment.
2. **Credit Spread (BofA Weekly):**
   Sourced from the Federal Reserve Bank of St. Louis, this variable measures the difference in yields between non-investment-grade corporate bonds and comparable Treasury bonds. It is an indicator of risk appetite in the broader credit market and may correlate with investor demand for CAT bonds.
3. **Expected Loss:**
   This variable represents the average notional amount expected to be lost on CAT bond sponsorships. It is a direct measure of the perceived risk embedded in CAT bonds, which influences yield levels.

It is also worth noting that the Total Coupon CAT Bond Market (USD) is defined as the sum of the Average CAT Bond Market Yield and the Money Market Rate, this metric provides a comprehensive measure of the total expected payout to investors. This is a result of investors receiving the returns on the collateral on top of interest from the bond, with the collateral being invested by the SPV in Money Market Funds.

### Trends and Observations

The average CAT bond yield exhibits variation over time, influenced by macroeconomic conditions, changes in risk perception, and market dynamics. All 3 are plotted from 2010 to the present in Figure 1. Next, figure 2 compares the movements of credit spread and the reinsurance market yield over time.

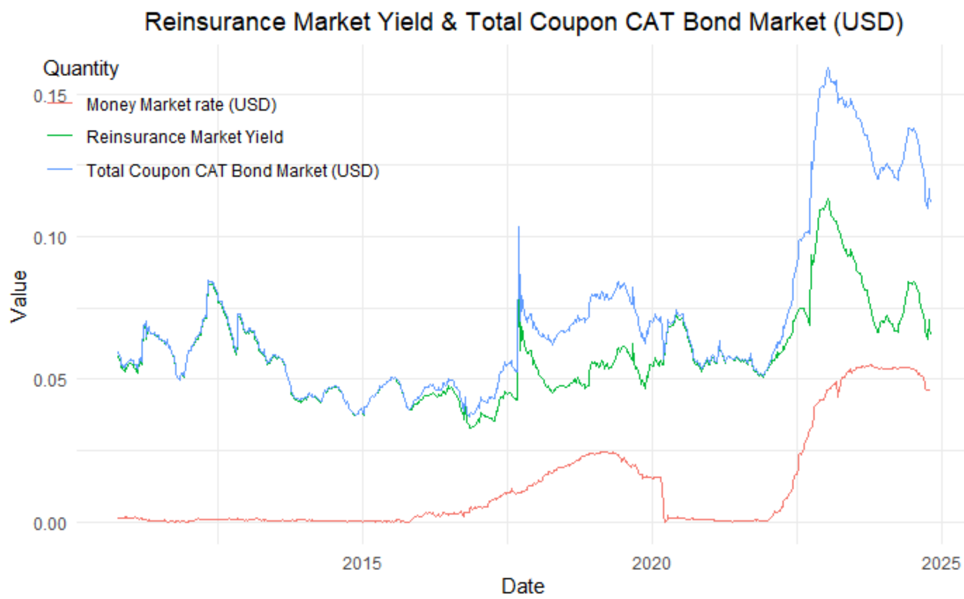


Figure 1. Reinsurance Market Yield & Total Coupon CAT Bond Market (USD)

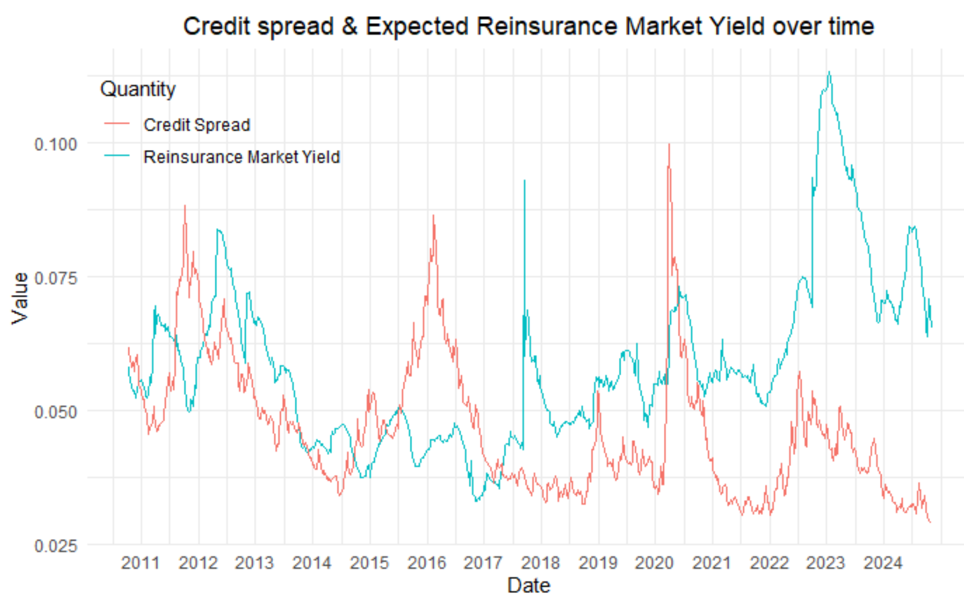


Figure 2. Credit Spread & Reinsurance Market Yield

The credit spread plot appears to be a lagged version of the reinsurance market yield. We kept this in mind when doing the correlation analysis. Next, we plotted Expected loss against reinsurance market yield to examine their relationship, shown in Figure 3.

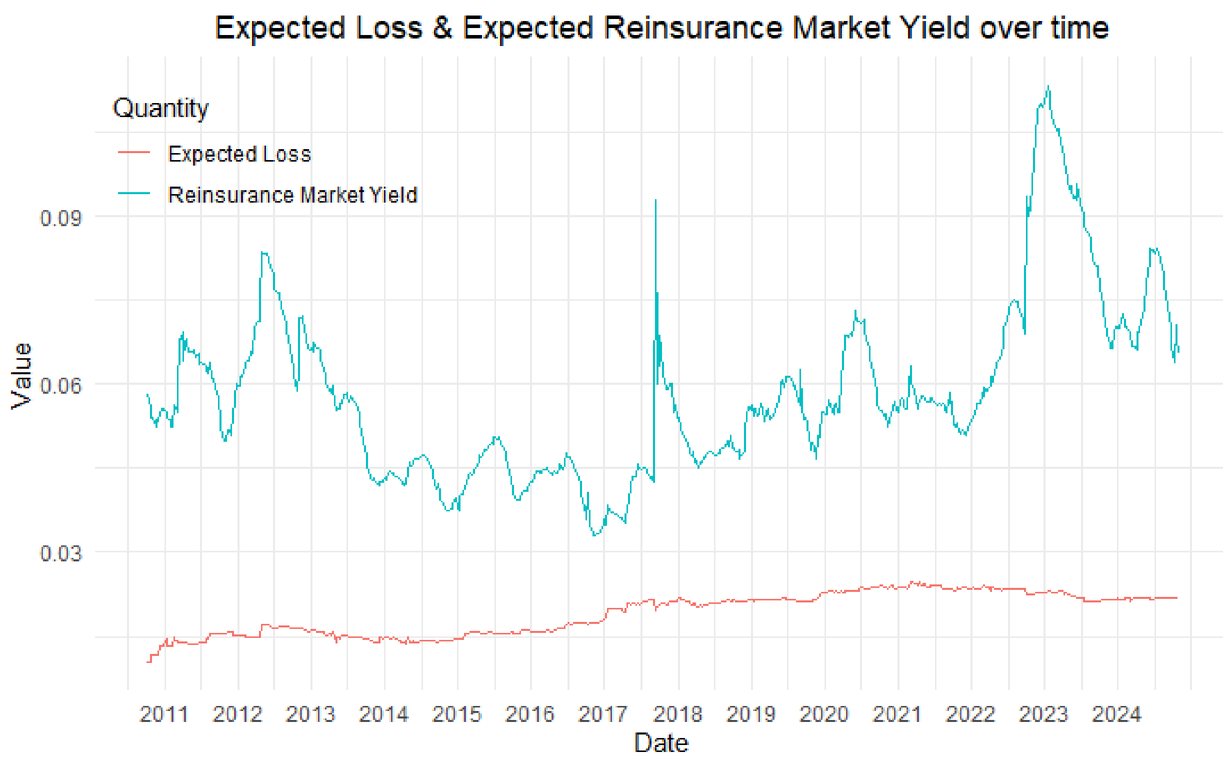


Figure 3. Expected Loss & Reinsurance Market Yield

For the expected loss variable, we observed that its values provide insight into the risk-reward tradeoff inherent in CAT bonds, with higher expected losses corresponding to higher yields as bonds with a greater probability of default need to have higher potential reward to offset this.

## Data Pre-processing

In the data preprocessing phase, we split the dataset into training and test sets. The training data spans from 2010-10-08 to 2020-10-08 (the first 10 years), while the test data covers 2020-10-08 to 2024-10-25 (the remaining period). We then examine the stationarity of only the training data to avoid lookahead bias. For the response variable (CAT bond yield), this helps determine how many times to difference the data when using ARIMA models. For the explanatory variables, we used cross-correlation plots to assess the correlation with the response, and ccf plots assume stationarity.

For each dataset, we examined the stationarity in the original data using a graphical method, i.e., the ACF graph, and we confirm whether the data is stationary after differencing (if necessary) by finding the characteristic roots. Below are the results we obtained:

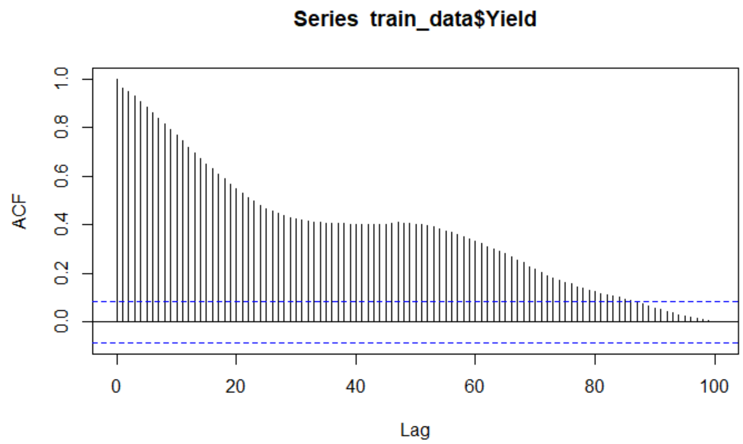


Figure 4. ACF of Reinsurance Market Yield Data

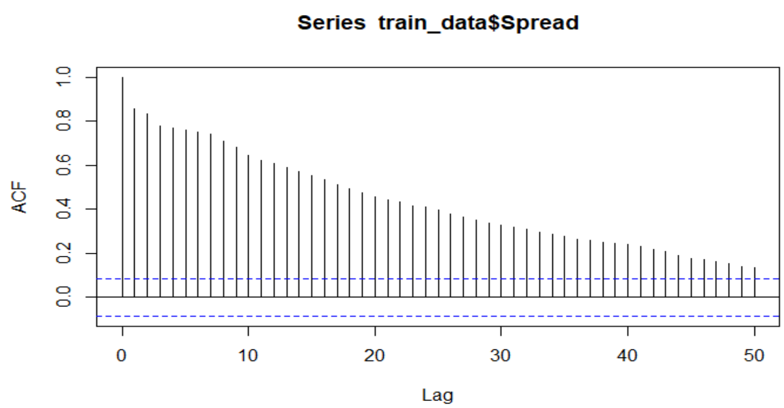


Figure 5. ACF of BofA Credit Spread Data

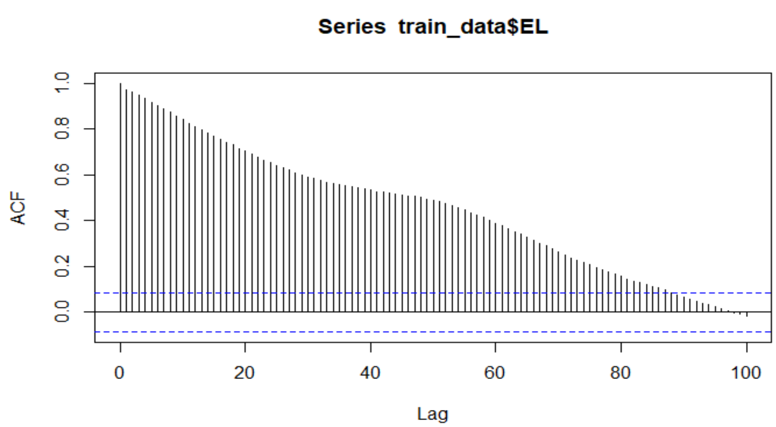


Figure 6. ACF of Expected Loss Training Data

All ACF graphs of the training data show correlations slowly decaying from 1, showing that the data is not stationary. Each series is differenced and then the ACF plotted again:

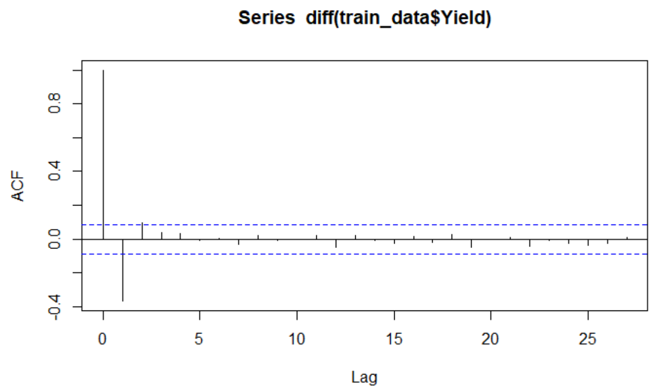


Figure 7. ACF of Differenced Reinsurance Market Yield Data

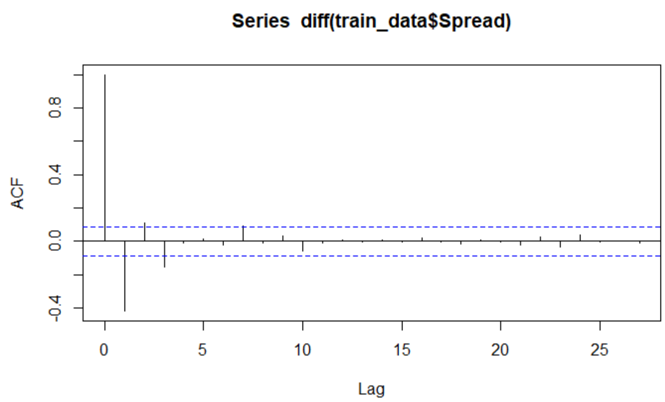


Figure 8. ACF of Differenced BofA Credit Spread Data

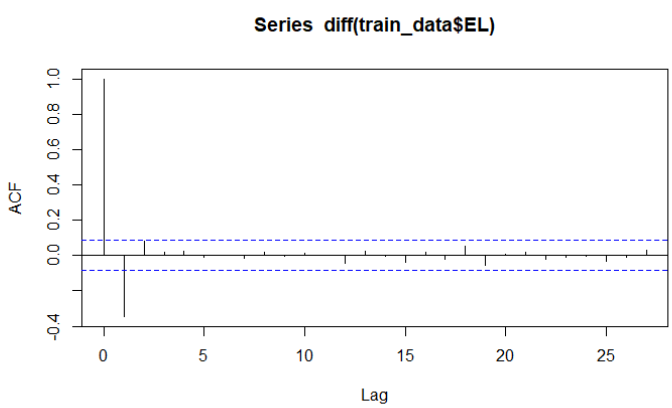


Figure 9. ACF of Differenced Expected Loss Data

These series only have a few significant lagged correlations. Next, the PACF plot each of the differenced series is plotted to examine the significant cross correlations (with alpha = .05):

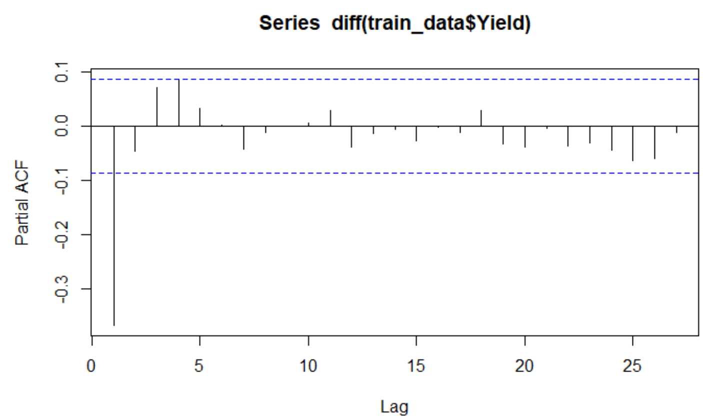


Figure 10. PACF of Differenced Reinsurance Market Yield Data

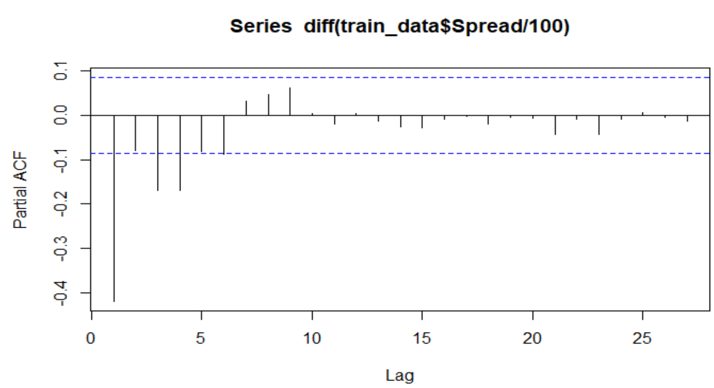


Figure 11. PACF of Differenced BofA Credit Spread Data

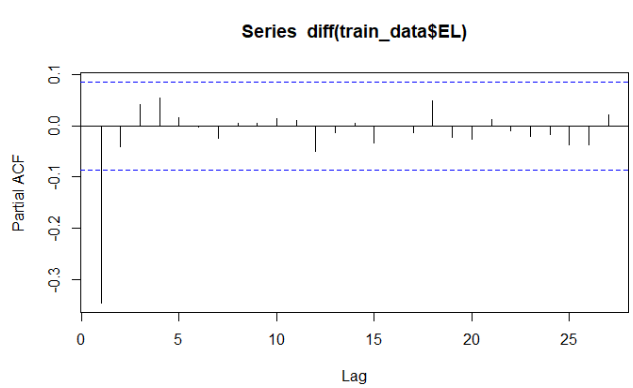


Figure 12. PACF of Differenced Expected Loss Data

For both the differenced Expected Loss and differenced Reinsurance Market Yield variables, there is only 1 statistically significant autocorrelation, at lag 1. Assuming a time series equation where only the previous term is considered, if we were to plug in the lag 1 correlation as $\alpha_1$ for both of these variables (with $\rho_{1,EL} = -.3465$ and $\rho_{1,mktYield} = -.367$, respectively), find $\theta_1(B)$, and solve for its roots, we would discover that both of these series are now stationary. However, for the differenced  Bank of America credit spread covariate, it is noteworthy that correlations with lags up to 5 were statistically significant in the PACF. Therefore, we used R's AR function to find the optimal AR process - based on AIC - considering lags up to 5. We used coefficients from that resulting model were used to generate $\theta_5(B)$, and then leveraged R's polyroot function to examine the roots of said equation. All roots were significantly greater than 1, indicating that this variable is stationary too. We could then proceed with cross-correlation analysis.

## Cross-Correlation Analysis

Then, we wanted to explore the cross-correlations between variables to see if there is any dependency between two time series variables at different time lags. First, we examine the cross-correlation between the differenced Credit Spread and CAT Bond Yield and plot the autocorrelation between these two variables at different lags.

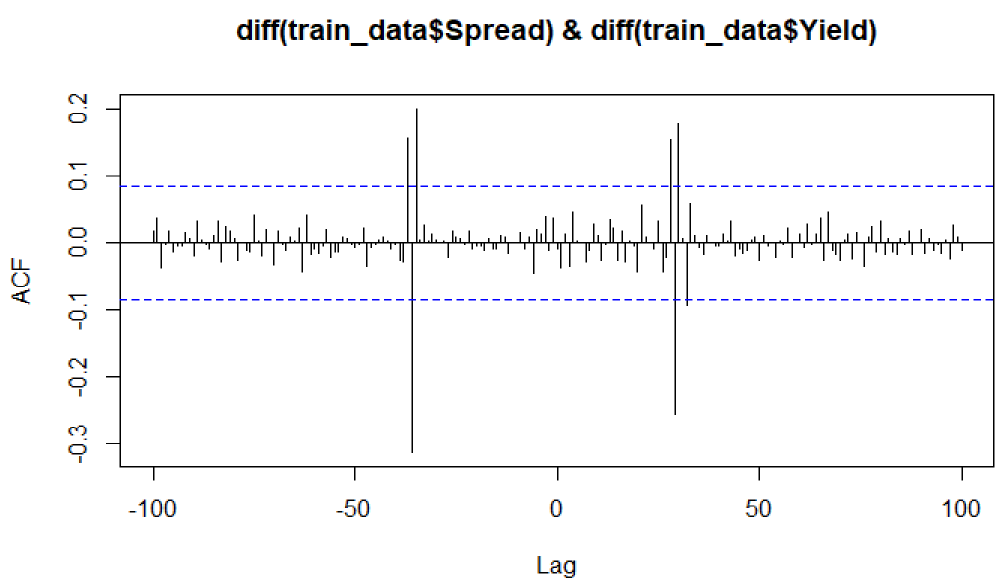


Figure 13. Cross Correlation between Differenced Credit Spread and CAT Bond Yield Data

From Figure 13, the cross-correlation plot shows that there is a significant autocorrelation at lag -36 (ACF value of -0.314), which suggests that the Credit Spread at lag -36 (i.e., about 36 periods ago) is significantly related to the change in the CAT Bond Yield. This suggests a negative relationship between the two variables at this specific lag, meaning that a change in the Credit Spread approximately 36 periods ago may have influenced the current change in CAT Bond Yield. Next, we examine the cross-correlation between the differenced Expected Loss and CAT Bond Yield data.

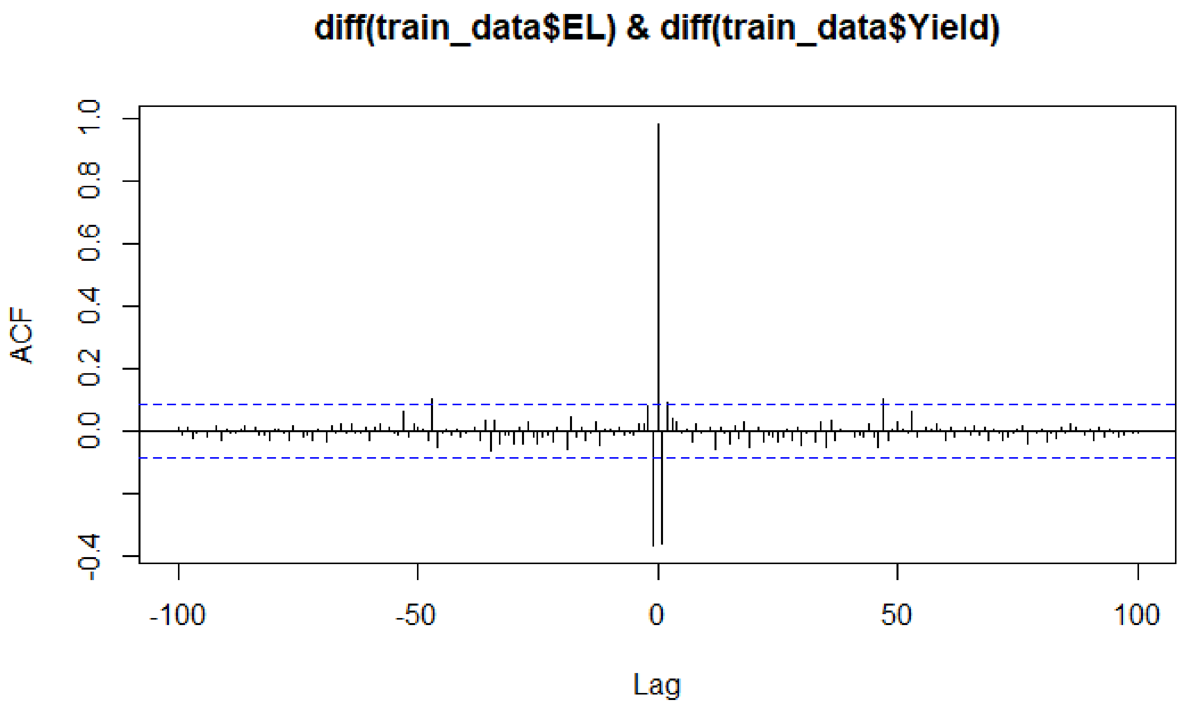


Figure 14. Cross-correlation between Differenced Expected Loss and CAT Bond Yield Data

The plot shows that the strongest correlation occurs at lag 0, meaning that the change in Expected Loss and the change in CAT Bond Yield are most strongly correlated in the current period (no lag). In other words, this cross-correlation analysis indicates a strong immediate relationship between the Expected Loss and CAT Bond Yield in the same period, implying that Expected Loss is a potentially important variable for predicting CAT Bond Yield in the same time frame for us to examine further. As a follow-up, we decided to create a scatter plot of Expected Loss and Reinsurance Market Yield, both undifferenced:

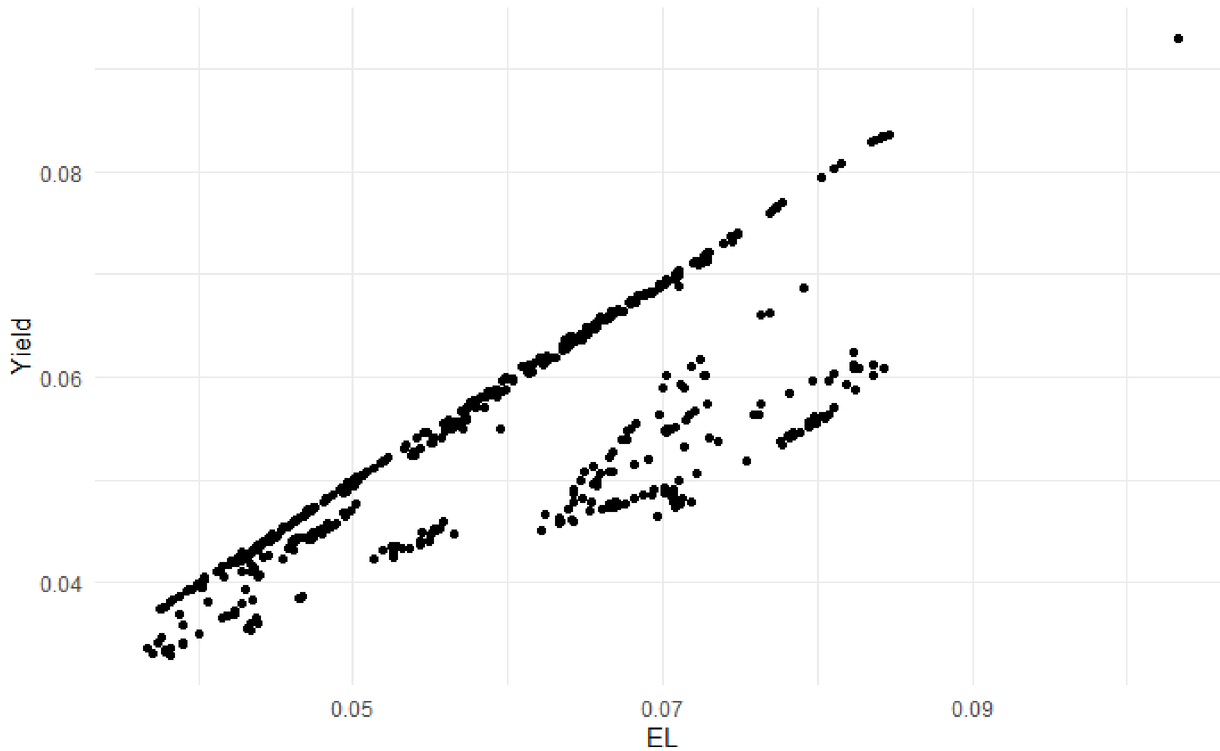


Figure 15. Relationship between Expected Loss and CAT Bond Yield

As the plot indicates, there is a strong correlation between Expected Loss and CAT Bond Yield. This led us to conclude that we could include expected loss as an exogenous covariate in regression without differencing.

In summation, we found that both Expected Loss and Credit Spread demonstrated predictive power in determining the expected CAT Bond Market Yield. Although the time series plot suggested that Credit Spread is leading Expected CAT Bond Yield, the strongest correlation observed was negative, which may indicate a potentially spurious relationship. This negative correlation calls for further investigation into the relationships between these variables and their influence on the CAT Bond Market Yield.

## 3. ARIMA Model Results & Comparisons

Finally, we developed three ARIMA models to predict CAT Bond Market Yield. The first model focused solely on autoregressive terms for market yield. The second model incorporated Expected Loss as an additional explanatory variable, aiming to improve prediction by considering potential risk factors. The third model further expanded by adding Credit Spread, examining its combined effect with Expected Loss on the market yield. These models were compared using BIC to evaluate their predictive performance.

### Model 1: ARIMA(1,1,0)

The first model we developed is the baseline model ARIMA(1,1,0) with no exogenous variables.

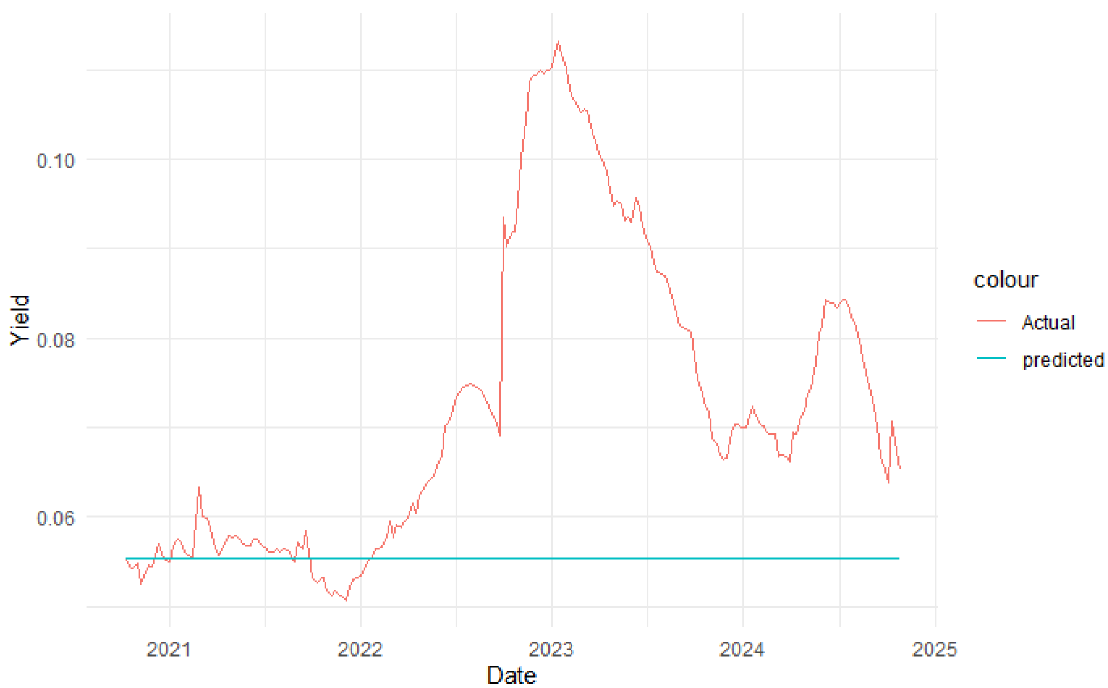


Figure 16. Baseline ARIMA(1,1,0) Model Yield vs. Date

From Figure 16, we see that in the baseline model ARIMA(1,1,0), the predicted CAT Bond Market Yield is represented by a straight line, indicating a simple trend with no significant fluctuations. The baseline model, which only uses autoregressive terms, does not fully capture the complexities of the data, and is only used as a baseline comparison with other models. It suggests that including additional covariates like Expected Loss and Credit Spread could improve the model's accuracy in predicting market yield. The AIC of the model is -4627.05.

### Model 2: ARIMA(1,1,0) + Expected Loss

The second model we developed is ARIMA(1,1,0) with Expected Loss as an exogenous variable.

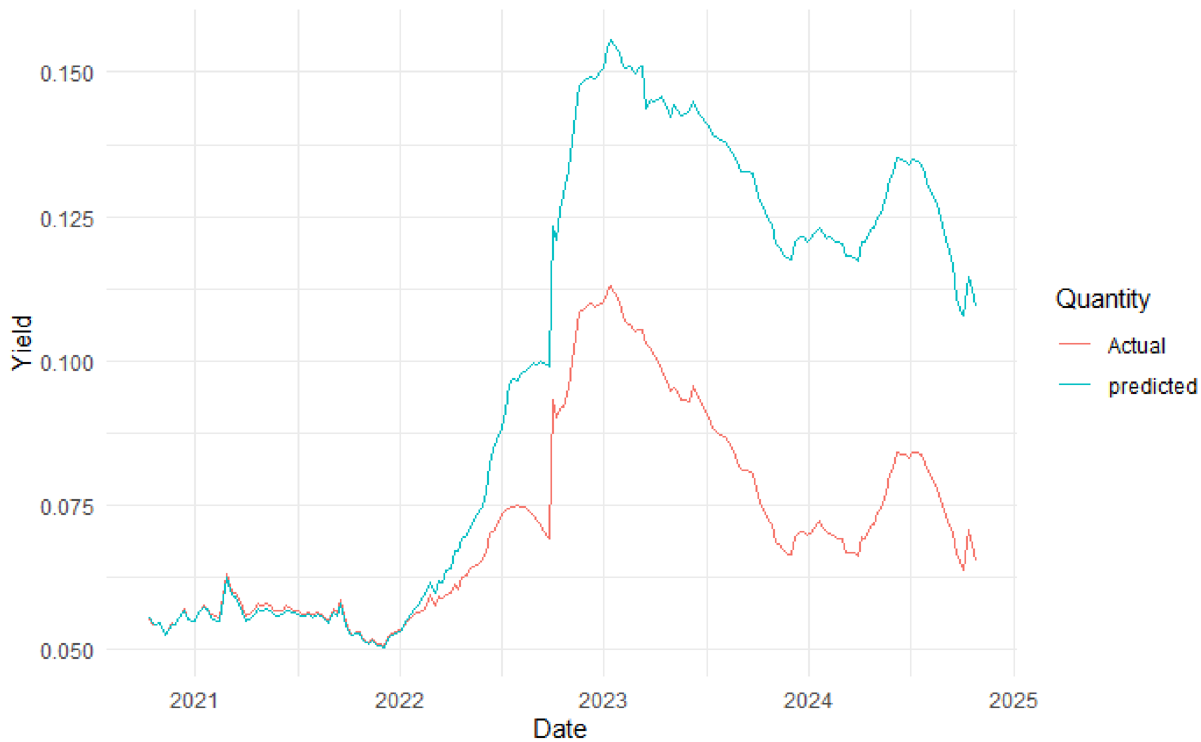


Figure 17. ARIMA(1,1,0) Model with EL as Exogenous Variable Yield vs. Date

The second model also used ARIMA(1,1,0), but with Expected Loss as an exogenous variable. The model's coefficients indicate a strong and statistically significant relationship between Expected Loss (coefficient = 0.9742, s.e. = 0.0061) and the Catastrophe Bond Market Yield. The log-likelihood of the model is 3232.36, and the AIC is -6458.72, suggesting a good fit. Figure 17 compares predicted yields to actual yields over time. The predicted yields follow the general trend and pattern of the actual yields but are consistently higher. This difference is most noticeable in recent years, where residuals increase. A potential reason for the increase in Expected Loss in recent years could be the rising frequency and severity of natural disasters driven by climate change. Events like hurricanes, wildfires, and floods have become more frequent and intense, leading to higher projected losses in the catastrophe bond market. This also suggests missing factors or market changes that the model does not account for, indicating room for improvement.

### Model 3: ARIMA(1,1,0) + Expected Loss + Credit Spread (lag 36)

The third model was similar to the second model, but with the addition of Credit spread at lag 36 as yet another exogenous variable.

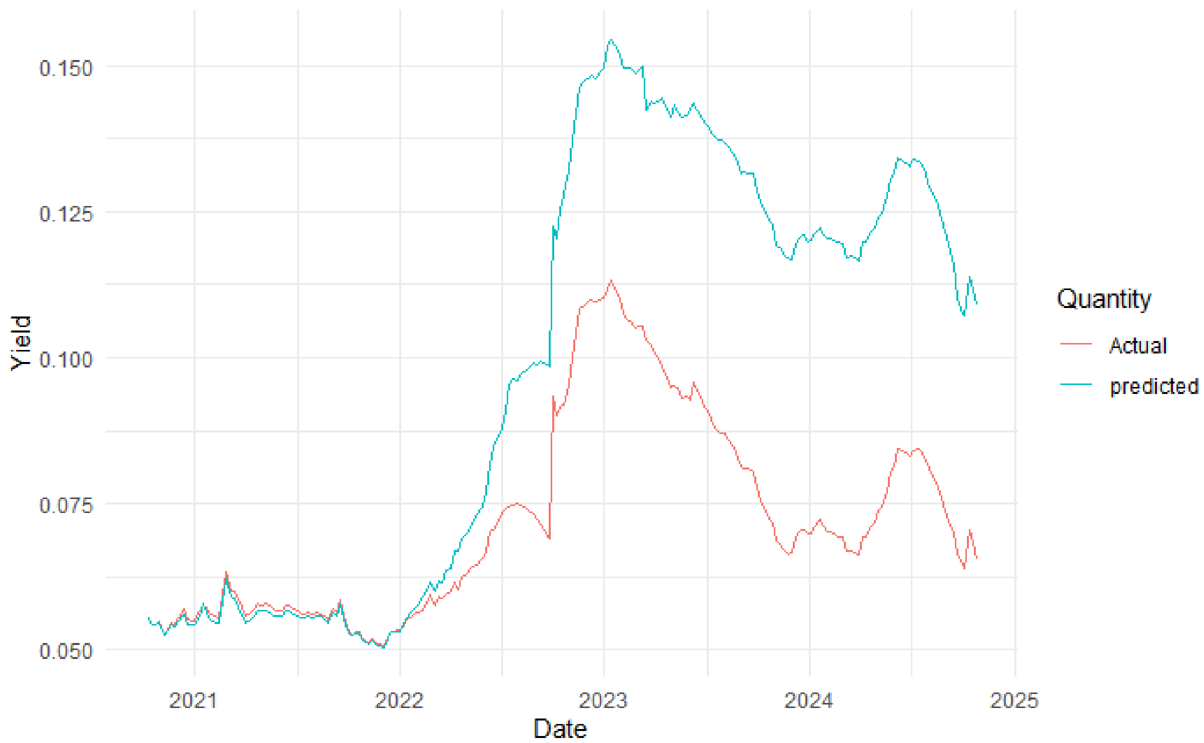


Figure 18. ARIMA(1,1,0) with expected loss and credit spread (lag 36)

Similar to the previous model, the predicted yield tends to be higher than the actual yield for most of the time, though both predicted and actual yields exhibit similar trends. This suggests that adding the credit spread does not significantly alter the overall pattern captured by the model. The inclusion of Credit Spread provides additional explanatory power, but the general behavior of the predicted and actual yields remains closely aligned. AIC for this model is -5985.47.

## Model Comparison

When comparing the three models, we chose BIC as the criterion for model selection due to the differences in dataset size, as the 1st 36 datapoints were removed from the training set due to the absence of a lag 36 credit spread value.. The BIC values indicate that Model 2 (ARIMA(1,1,0) with Expected Loss) is the best-fitting model. Although the inclusion of Credit Spread in Model 3 does not significantly impact the model, it does slightly improve the predictions on the test set. However, the BIC still favors Model 2, suggesting that adding Credit Spread does not provide substantial additional predictive power.

| model <chr> | BIC <dbl> |
|---|---|
| m1 | -4624.794 |
| m2 | -6452.200 |
| m3 | -5974.908 |

Figure 19. Model BIC Values Comparison

We then generated a graph comparing the Model Predictions of the Reinsurance Market Yield with the Actuals (Figure 20). Both Model 2 and Model 3 follow the pattern of the actual data closely. However, around mid-2022, the residuals (the difference between predicted and actual values) begin to increase. This suggests that the models may not fully capture certain factors influencing the yield, indicating the potential presence of a missing covariate. Further exploration of additional variables could help improve model accuracy and reduce the increasing residuals observed in the later part of the data.

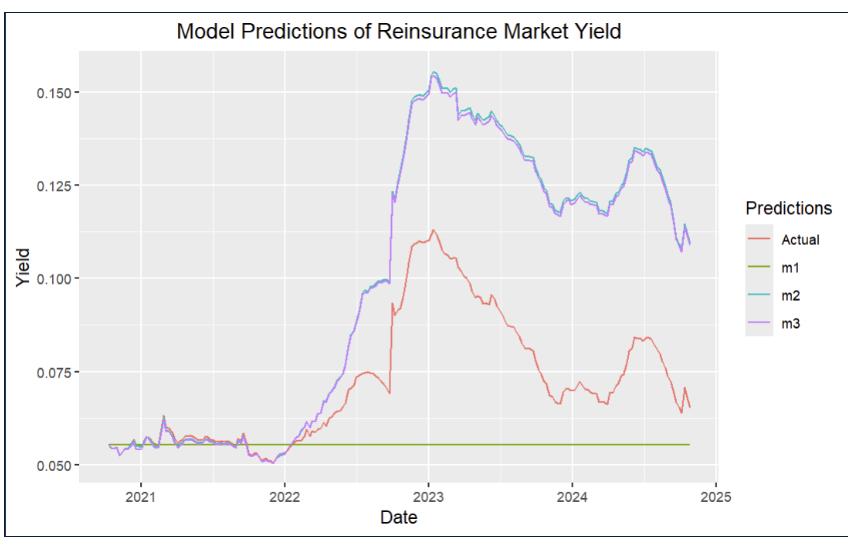


Figure 20. Model Predictions of Reinsurance Market Yield

## 4. Conclusion

This study aimed to study if we can predict the market yield of catastrophe bonds (CAT bonds) using Expected Loss and Credit Spread as key covariates. We developed three ARIMA models, adding autoregressive terms, Expected Loss, and Credit Spread to evaluate their predictive power. Both Expected Loss and Credit Spread showed predictive value, with Expected Loss being the stronger predictor. The inclusion of Credit Spread slightly improved predictions, but its impact was less significant. Although Credit Spread and Expected CAT Bond Yield were negatively correlated, this may be spurious. Overall, the models with these covariates tracked market yield patterns well.

Despite these findings, there were limitations. The analysis only used the most significant lag of Credit Spread, leaving room for further improvement. We also lacked data on total returns rather than just Expected Yield, which made it harder to compare CAT bond returns with general market performance. Finally, there may be missing covariates, such as inflation or CAT activity, that could explain the decrease in prediction accuracy in 2022. These factors suggest that the models could be refined in future work.

## 5. References

Artemis. (n.d.). *Catastrophe bond market yield over the past 10 years*. Retrieved from https://www.artemis.bm/catastrophe-bond-market-yield/

American Academy of Actuaries. (2022, June 14). *What is a CAT bond/ILS?* Retrieved from https://www.actuary.org/sites/default/files/2022-06/ILS_20220614.pdf

Insurance Journal. (2024, June 24). *Why is demand increasing?* Retrieved from https://www.insurancejournal.com/news/international/2024/06/24/780786.htm

Swiss Re. (2014, August). *Index methodology: CAT bond indices methodology*. Retrieved from https://www.swissre.com/dam/jcr:307452ca-9664-4772-96f9-7c11f80109b2/2014_08_ils_cat_bond_indices_methodology.pdf

Swiss Re. (2024, February). *Performance of CAT bonds: ILS market insights*. Retrieved from https://www.swissre.com/dam/jcr:bb189e59-a15f-49df-a250-07b2c6b2d9bd/2024-02-sr-ILS-market-insights-feb-2024.pdf