QF604 - Econometrics of Financial Markets

Group Project

Predicting Natural Gas Futures' Volatility using Climate Risks

Finance Research Letters 55 (2023) 103915

Group Members:

Bryan Wilson

Durgheshwari Elamurugu

Huang Ziyan

Sharon Gautama

Sugintha Sivaneswaran

Summary

The study "**Predicting natural gas futures' volatility using climate risks**" explores the impact of climate risks, specifically climate policy uncertainty (CPU) and climate-related disasters, on the price volatility of natural gas futures.

Climate change poses significant risks to financial markets, including physical and transition risks, leading to economic losses and market instability:

- Physical Risks (extreme weather events and natural disasters) indicated by climate-related disasters.
- Transition Risks (uncertainties related to the shift towards a low-carbon economy) indicated by climate policy uncertainty (CPU)

The energy futures market is closely linked to climate change, making it susceptible to climate risks. Natural gas is chosen as a proxy for energy futures due to its significance. So, this study assesses how CPU and climate-related disasters impact natural gas futures price volatility using the GARCH-MIDAS model.

Methodology

The study's methodology section outlines the GARCH-MIDAS (GM) model to efficiently incorporate daily natural gas futures prices with monthly CPU indices and disaster frequencies. The MIDAS (Mixed Data Sampling) approach allows for incorporating variables sampled at different frequencies into the model. For instance, a GARCH-MIDAS model can use monthly macroeconomic indicators to model daily stock returns volatility, capturing the impact of slow-moving variables on fast-moving financial data.

This combines a daily GARCH process for short-term volatility with a MIDAS polynomial for long-term trends, leveraging lower-frequency variables for better predictions of high-frequency variables. The model's structure involves a mean reverting unit daily GARCH process and a MIDAS polynomial to represent the short and long-term components of conditional volatility, respectively.

In addition to the GARCH-MIDAS method used in the paper, we explored using another type of model, which is the Double Asymmetric GARCH-MIDAS (DAGM) model. This model introduces additional parameters to further address the leverage effect, a common phenomenon in financial markets where negative returns increase future volatility more than positive returns of the same magnitude.

The Double Asymmetric GARCH-MIDAS (DAGM) model distinguishes itself from the traditional GARCH-MIDAS by addressing asymmetry in both its components: the high-frequency data, such as daily financial returns, and the low-frequency data, such as macroeconomic indicators. This dual approach to incorporating asymmetry allows the model to reflect the nuanced impacts of the economic indicators more accurately on volatility, enhancing its predictive power by acknowledging that both positive and negative changes in these two types of data can have different effects on financial market volatility. The 'RUGARCH' package will be used for GARCH, and the 'RUMIDAS' package will be used for GM and DAGM models. The primary language used is R.

Data Description

The dataset consists of:

- daily NYMEX natural gas futures prices
- monthly US Climate Policy Uncertainty (CPU) indices
- monthly frequencies of climate-related disasters.

Sample Period: The data spans from 01/01/1991 to 29/07/2022 representing the longest available period.

Replication of Empirical Findings

As part of the project's requirement, we were able to successfully replicate and graphically illustrate the empirical findings to a certain extent of the paper in the following 3 aspects.

Original Paper Graph Results

- Daily Return of NYMEX Natural Gas Futures:
    - Shows the daily price return of natural gas futures from 1991 to 2022.
    - The returns fluctuate around zero, with some spikes indicating days of high price volatility.
- Monthly Frequency of Natural Disasters:
    - Illustrates the number of natural disasters occurring each month from January 1991 to July 2022.
    - There are peaks indicating months with higher frequencies of natural disasters.
- U.S. Climate Policy Uncertainty Index (CPU):
    - Depicts the level of uncertainty in U.S. climate policy from January 1991 to June 2022.
    - The index increases over time, with a noticeable upward trend in recent years, suggesting growing uncertainty.

These graphs are used to analyze the potential impact of natural disasters and policy uncertainty on the volatility of natural gas futures. We are also able to understand from **Figure** 1 that there is a **presence of unit root**.

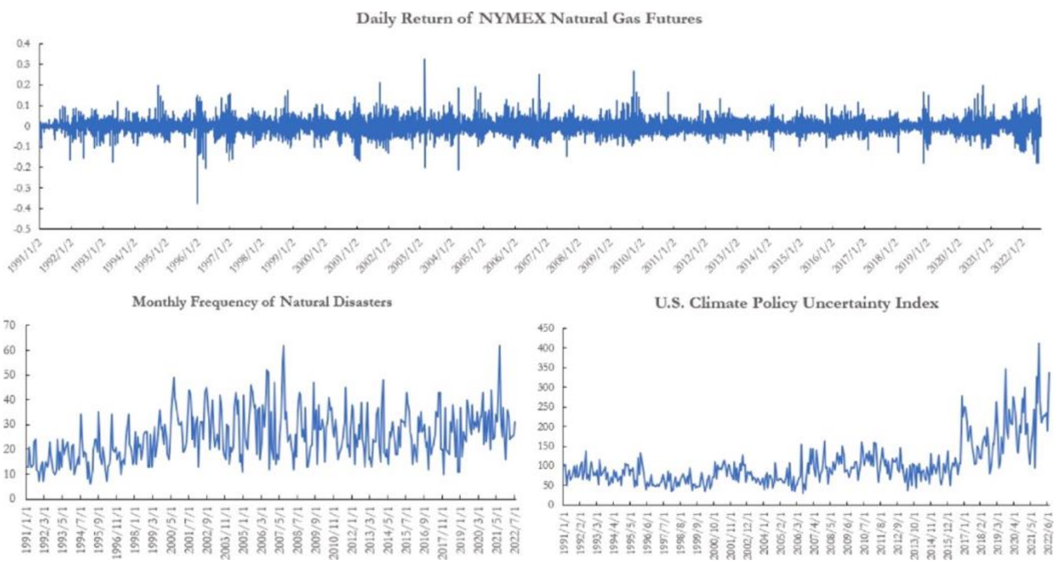


Figure 1 Trends of the daily return of NYMEX natural gas futures and two climate-related predictors

Original Paper Statistics

Table 1 Insight of Descriptive Statistics

Table 1
Descriptive statistics.

Variable	Obs.	Mean	Std.	Min	Max	Skewness	Kurtosis	ADF
Natural gas futures return	8021	0.000	0.035	−0.376	0.324	0.103	9.555	−18.117***
CPU Index	379	102.536	57.212	28.162	411.289	1.929	7.587	0.678
d. CPU Index	379	379	0.004	0.371	−1.701	−0.223	3.787	−18.132***
Natural disasters frequency	379	25.311	9.972	6	62	0.600	3.178	−2.6280*
d. Natural disasters frequency	379	0.002	0.437	−1.447	1.340	−0.027	3.319	−18.014***

Notes: ADF is the t-statistics for the augmented Dickey–Fuller test. The symbol "d." denotes the first-order difference after log-transformation. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

From **Table 1**, we can draw the following conclusions. The table indicates that the natural gas futures return has a mean of zero and exhibits leptokurtic behavior (high kurtosis). The CPU Index shows positive skewness and high variability. The first-order differences of the CPU Index and natural disasters frequency are used to achieve stationarity, as indicated by the significant negative ADF values, suggesting that the differenced series are stationary.

In the replication process, we met with the limitation of not being able to access dataset for the years 1991 through 1993. However, we were able to generate the overall shape and distribution like the original paper.

Replication Graphic Results

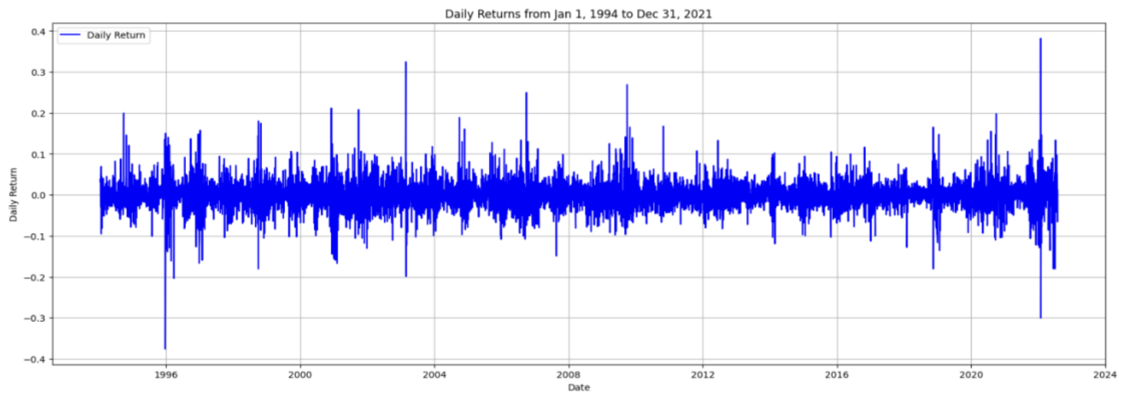


Figure 2 Replication Dataset of Daily Return of NYMEX Natural Gas Futures

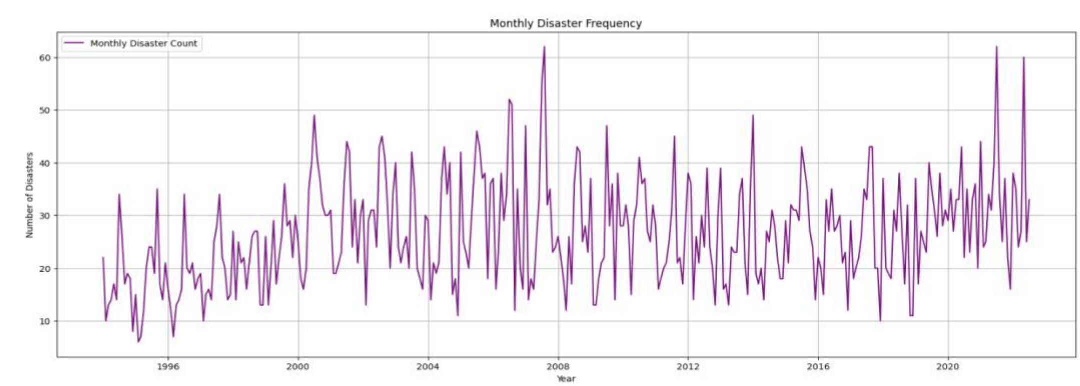


Figure 3 Replication Dataset of Monthly Frequency of Natural Disasters

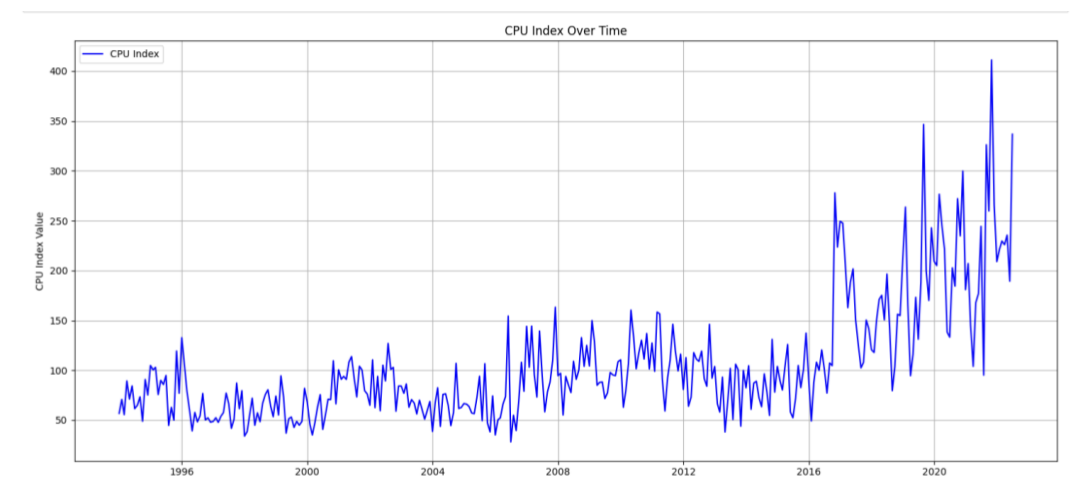


Figure 4 Replication Dataset of U.S. Climate Policy Uncertainty Index (CPU)

Although the results are not shown in the original paper, as mentioned above there is a presence of unit root. Hence, by performing 1st differencing after the log-transforming the results, we were able to render the results stationary for the 2 factors - CPU index and Disaster Frequency as shown in **Figure 5** and **Figure 6**.The purpose of log-transforming is to reduce the impact of outliers by compressing the data's range and bringing extreme values closer to the mean. This is because skewed data can make it challenging to interpret results and fit models leading to inaccuracy.

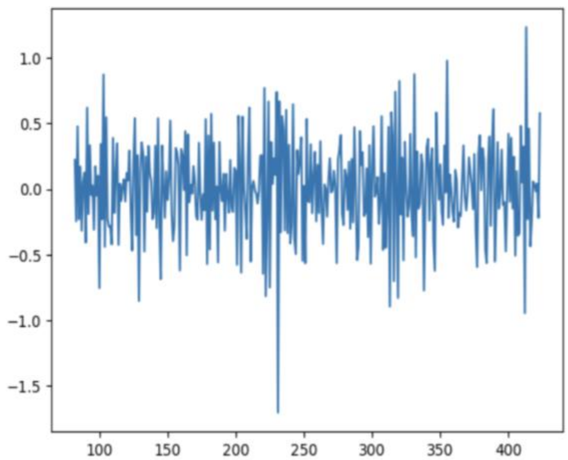


Figure 5 Stationary Test of CPU Index

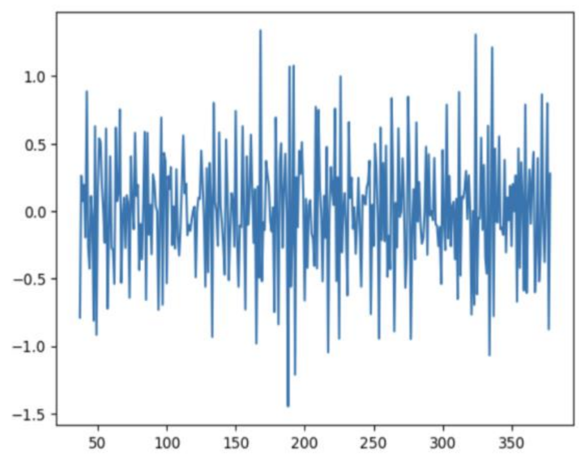


Figure 6 Stationary Test of Disaster Index

Replication Statistics

Figure 8 represents the test results represents the test results before after log transformation and 1st - differencing. We focus on the Augmented Dickey-Fuller Test (ADF) on the below. Following the same logic also applies to the CPU index test results after log transformation and 1st - differencing**.** The ADF test results give us a test statistic and p-value. The test statistic is compared to the critical values at different significance levels of 1%, 5%, and 10%.

Test result interpretation

- If the test statistic is more negative (less negative) than the critical value, you can (cannot) reject the null hypothesis and conclude that the time series is stationary (not stationary & has unit root).
- If the p-value is less than (greater than) the significance level, you can (cannot) reject the null hypothesis and conclude that the time series is stationary (not stationary & has unit root).

Table 2 Insight of Descriptive Statistics of Original Paper

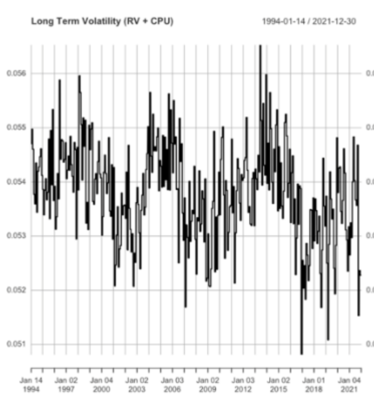
Descriptive Statistics								
Variable	**Obs**	**Mean**	**Std.**	**Min**	**Max**	**Skewness**	**Kurtosis**	**ADF**
Natural gas Futures Return	7169	0.000	0.0362	-0.376	0.382	0.253	8.338	-17.428
CPU Index	343	104.623	59.439	28.162	411.289	1.813	3.991	0.439
d. CPU Index	343	0.005	0.372	-1.701	1.233	-0.226	0.923	-7.569
Natural Disasters Frequency	343	26.446	10.012	6	62	0.605	0.332	-2.7706
d. Natural Disasters Frequency	343	0.001	0.447	-1.447	1.34	-0.0309	0.329	-8.2489

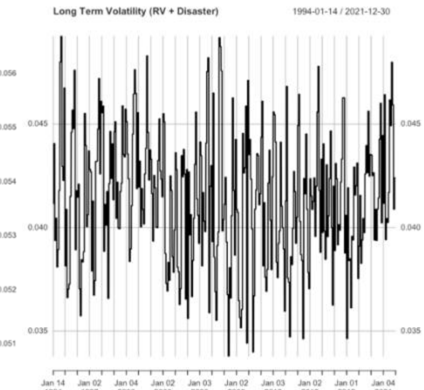
Figure 8 Key Results

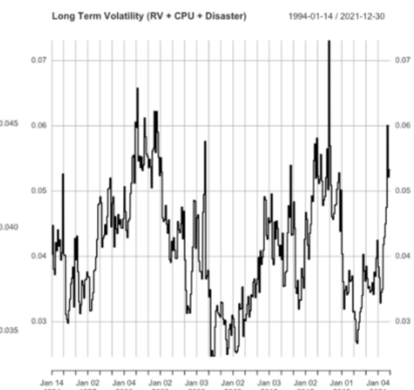
	Disaster frequency before differencing	**Disaster Frequency after log-transformation & 1st Differencing.**	**CPU Index before differencing**	**CPU Index after log-transformation & 1st Differencing.**
ADF stat	-2.770696	-8.248828	0.438569	-7.568764
p-value	0.062583	5.460646e-13	0.982881	2.8784425e-11
lags	14	13	12	11
Obs.	327	328	329	330
1% signif	-3.450507	3.450445	-3.450383	-3.450322
5% signif	-2.870419	-2.870392	-2.870365	-2.870338
10% signif	-2.571500	-2.571486	-2.571471	-2.57145

Let us take an example of the disaster frequency to interpret the result. With reference to **Figure 8**, the test statistic of the ADF is less negative than critical values at 1% and 5%. Hence, we cannot reject the null hypothesis and conclude that the times series of this dataset is not stationary. However, after the necessary actions have been performed such as log transformation and 1st – differencing. From **Figure 8**, we can see that the test statistic of ADF is more negative than the critical values at 1%, 5% and 10%. Therefore, we can reject the null hypothesis and conclude that we have rendered the dataset stationary. The same logic applies for the CPU index dataset, and we arrive at the same conclusion of stationarity as referenced in **Figure 8**.

Secular Components







The three graphs illustrate distinct volatility patterns, each reflecting the unique influence of the variables included in the models. The graph for the RV + CPU + DISASTER model appears to represent a confluence of patterns observed in the RV + CPU and RV + DISASTER models. This suggests that the combined model is correctly capturing the individual effects of CPU and disaster factors on long-term volatility, as its pattern is a composite of the other two models.

Replication Comparison

With the replication of the GARCH-MIDAS model, a flexible framework that integrates daily natural gas futures prices with low-frequency predictors such as climate policy uncertainty (CPU) and climate-related disaster frequencies, was employed. Four distinct models are considered: the basic RV model, the RV + CPU model, the RV + Disaster model, and the comprehensive RV + CPU + Disaster model. The replication process yielded results that closely matched the original findings in several aspects.

The GARCH(1,1) model is as follows,

$$g_{i,t} = (1 - \alpha - \beta) + \alpha \frac{\left(r_{i-1,t} - \mu\right)^2}{\tau_t} + \beta g_{i-1,t}$$

$$r_{i,t} = \mu + \sqrt{\tau_t g_{i,t}} \varepsilon_{i,t}, \forall i = 1,2, \cdots, N_t$$

Where,

$r_{i,t}$ is the daily return for the i-th day $(i = 1, \dots, N_t)$ of the period t

τ_t is the long-run component, varying each period t

$g_{i,t}$ is the short-run term, varying each day i of the period t

$\varepsilon_{i,t}$ is an iid error term which has a zero mean and unit variance.

In the tables, *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1-3 Replication Result of GARCH Model with basic RV

	RV		
	Report	**Replication**	**Difference**
μ	0.0004	0.0004	6%
ω	1.0000	1.0000	0%
α	0.0922***	0.0923***	0%
β	0.8921***	0.8992***	1%
LLF	14495.1348	14230.73	2%

In the basic RV model, we implied the GARCH (1,1) model with the only variable, the realized volatility of daily returns (RV). ω is the weighting scheme with a Beta lag structure, here must be 1. The difference between the original result and ours are quite small.

The short–run component of the GARCH–MIDAS model is,

$$g_{i,t} = (1 - \alpha - \gamma/2 - \beta) + \left(\alpha + \gamma \cdot I_{(r_{i-1,t} < 0)}\right) \frac{\left(r_{i-1,t}\right)^2}{\tau_t} + \beta g_{i-1,t}$$

And the long-run component is,

$$\tau_t = \exp\left\{m + \theta \sum_{j=1}^{K} \delta_j(\omega) X_{t-j}\right\}$$

Table 1-2 Replication Result of GARCH-MIDAS Model with RV and CPU

	RV+CPU		
	Report	**Replication**	**Difference**
α	0.0896 ***	0.0882***	-2%
β	0.8936 ***	0.903 ***	1%
m	-2.2895	-6.2604	173%
θ	0.3051***	0.3077***	1%
ω	35.7809 **	35.7809 **	0%
LLF	14,499.4019	14,231.5376	-2%
BIC	-28,927.8317	-28,418.7898	-2%
Lag	(36,36)	(36,36)	

Table 1-3 Replication Result of GARCH-MIDAS Model with RV and Disaster

	RV + Disaster		
	Report	**Replication**	**Difference**
α	0.0941 ***	0.0843 ***	-10%
β	0.8846 ***	0.9070 ***	3%
m	-2.3571	-6.4335	173%
θ	9.7665 ***	9.8037 ***	0%
ω	1.0000 ***	1.0011 ***	0%
LLF	14,640.5884	14,249.8749	-3%
BIC	-29,210.1354	-28,437.7503	-3%
Lag	(33,33)	(33,33)	

In the RV + CPU and RV + Disaster models, most of the parameters, including alpha, beta, theta, and omega, closely aligned with the original results, with errors ranging from -10% to 3%. Only the parameter $\boldsymbol{m}$ displayed a substantial difference of 173%, however it is statistically insignificant in both cases. Back to formula of GARCH-MIDAS above, the possible reason of the difference is the choice of database of these 2 variables.

Table 1-4 Replication Result of Double Asymmetric GARCH-MIDAS Model with RV, CPU and Disaster

RV+CPU+ Disaster			
Report		**Replication**	
$\boldsymbol{\alpha}$	0.0867 ***	$\boldsymbol{\alpha}$	0.0795 ***
$\boldsymbol{\beta}$	0.8926 ***	$\boldsymbol{\beta}$	0.9111 ***
m	-2.3571	**m**	-4.3661
$\boldsymbol{\theta_{RV}}$	0.7724	$\boldsymbol{\theta_{RV,CPU_{Positive}}}$	1.4378
$\boldsymbol{\theta_{CPU}}$	0.4676 **	$\boldsymbol{\theta_{RV,CPU_{Negative}}}$	-2.6967
$\boldsymbol{\theta_{Disaster}}$	9.5393***	$\boldsymbol{\theta_{RV,Disaster_{Positive}}}$	-0.3580
		$\boldsymbol{\theta_{RV,Disaster_{Negative}}}$	15.4299
$\boldsymbol{\omega_{RV,2}}$	1.0027	$\boldsymbol{\omega_{RV,CPU_{Positive}}}$	15.8211
$\boldsymbol{\omega_{CPU,2}}$	24.6131 **	$\boldsymbol{\omega_{RV,CPU_{Negative}}}$	2.3948 ***
$\boldsymbol{\omega_{Disaters,\ 2}}$	1.0000 ***	$\boldsymbol{\omega_{RV,Disaster_{Positive}}}$	21.2811
		$\boldsymbol{\omega_{RV,Disaster_{Negative}}}$	1.1733 ***
LLF	14,645.2893	**LLF**	14,258.8020
BIC	-29,201.7769	**BIC**	-28,419.1761
Lag	(33,33,33)	**Lag**	(33,33,33)

In the Double Asymmetric GARCH-MIDAS model, the long-run component has been changed to

$$\tau_t = \exp\left(m + \theta^+ \sum_{k=1}^{K} \delta_k(\omega)^+ X_{t-k} I_{(X_{t-k} \geq 0)} + \theta^- \sum_{k=1}^{K} \delta_k(\omega)^- X_{t-k} I_{(X_{t-k} < 0)}\right)$$

Here, due to the choice of different GARCH-MIDAS models, the different parameters have been yielded, which caused the difficulty in comparison with horizontal levels. However, it still exhibited similar trends, with most parameters closely matching the original results.

Summary

It can be observed that the ARCH terms α (alpha) and GARCH terms β (beta) in all the models were all statistically significant, confirming the existence of volatility clustering and persistence in natural gas futures price.

Furthermore, the sum of α and β is less than 1. This finding is significant as it indicates that the volatility shocks—sudden, unexpected changes in the volatility—are transitory rather than permanent.

The implication of α + β being less than 1 is that the impact of a shock to volatility is not indefinitely sustained. Instead, these shocks tend to diminish over time, allowing the volatility to revert to a long-term mean. This mean-reverting behavior of volatility is a critical aspect of financial time series, reflecting the inherent tendency of markets to stabilize following periods of disequilibrium.

The decay in the effect of volatility shocks is essential for the model to remain "well-behaved," meaning that it maintains stability over time and does not project escalating variance that diverges. This stability allows for the GARCH model to produce reliable forecasts for future volatility. Therefore, the condition of $\alpha + \beta < 1$ not only confirms the appropriateness of the models used but also underscores their capability to generate credible predictions.

The positive MIDAS slope coefficients, θ_{CPU} and $\theta_{Disaster}$, as depicted in Tables 12 and 13, suggest a direct correlation between CPU changes and disaster frequency with the long-term volatility of natural gas futures. Evolving climate policies and the increased occurrence of disasters are linked to heightened volatility in natural gas pricing. This relationship persists even when CPU, disaster frequency, and RV are collectively included in the tri-variate GARCH-MIDAS model, as shown in Table 14.

Furthermore, the larger log-likelihood function (LLF) values associated with all three multi-predictor models, compared to the basic RV model, indicate that the inclusion of CPU and disaster frequency provides additional insights into the price volatility of natural gas futures. This enhancement in LLF values underscores the additive predictive power of these variables, reinforcing the comprehensive nature of the model in capturing the dynamics of market volatility.

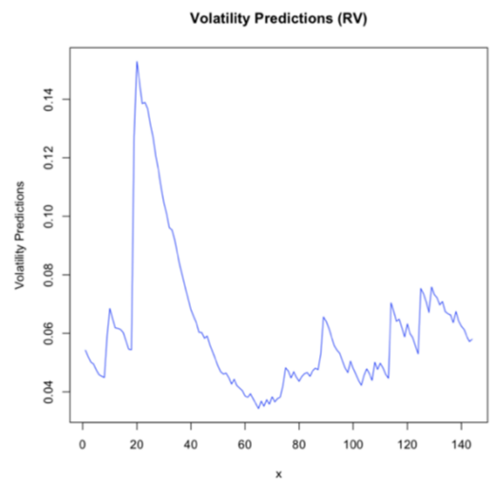
Additional Model in Replication

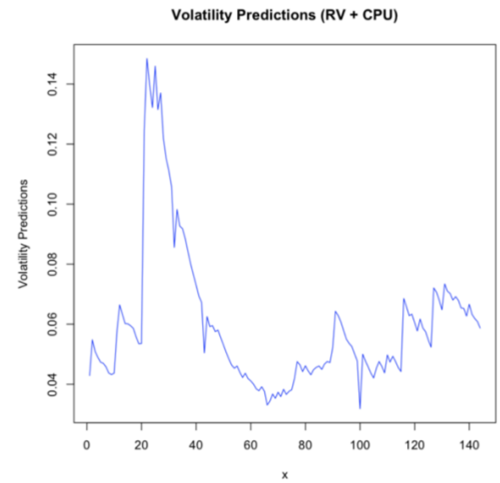
The extension of GARCH-MIDAS model in **Table 4**, which had been used in the tri-variate model above, is to apply Double Asymmetric GARCH-MIDAS on the other 2 datasets. Although they had been observed much more different results, they still have the meaning of reference in econometrical findings.

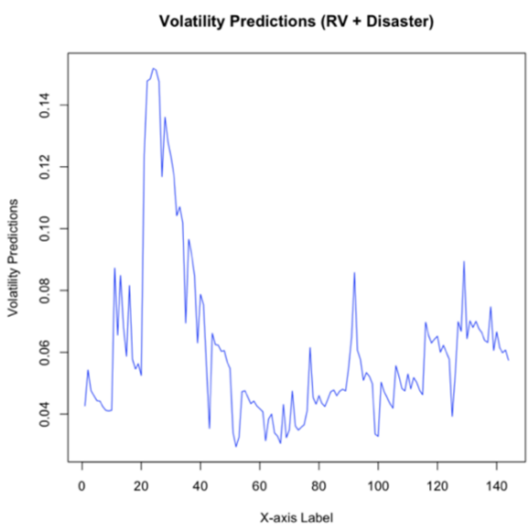
Table 4 Application of DAGM Model on RV+CPU & RV+Disaster

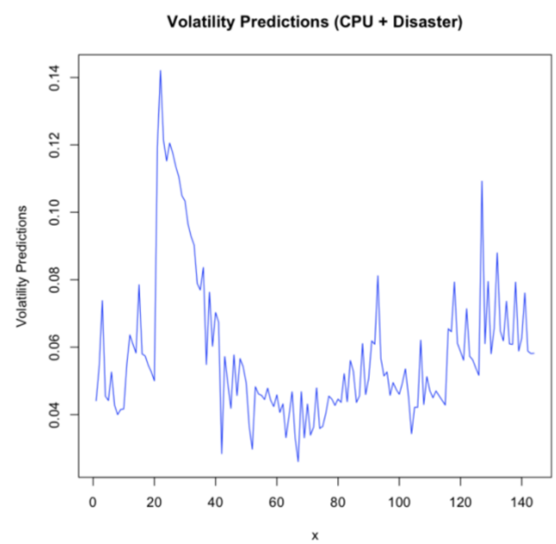
	RV+CPU	**RV + Disaster**
$\boldsymbol{\alpha}$	0.0882***	0.0843***
$\boldsymbol{\beta}$	0.903***	0.907***
m	-6.5356***	-3.7825***
$\boldsymbol{\theta_{RV,CPU_{Positive}}}$	1.1832***	-0.5235*
$\boldsymbol{\theta_{RV,CPU_{Negative}}}$	-0.5608*	14.6688***
$\boldsymbol{\omega_{RV,CPU_{Positive}}}$	18.9603***	13.163***
$\boldsymbol{\omega_{RV,CPU_{Negative}}}$	9.5365***	1.206***

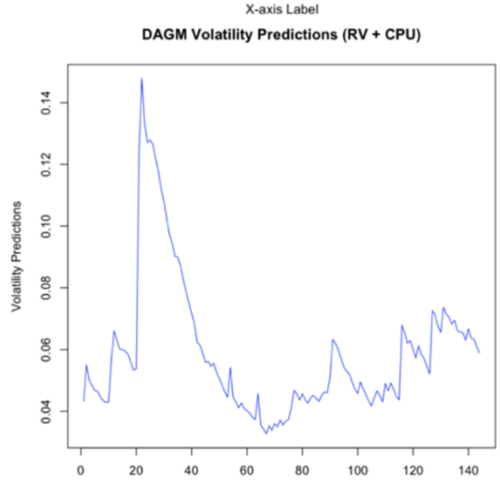
GRAPHS

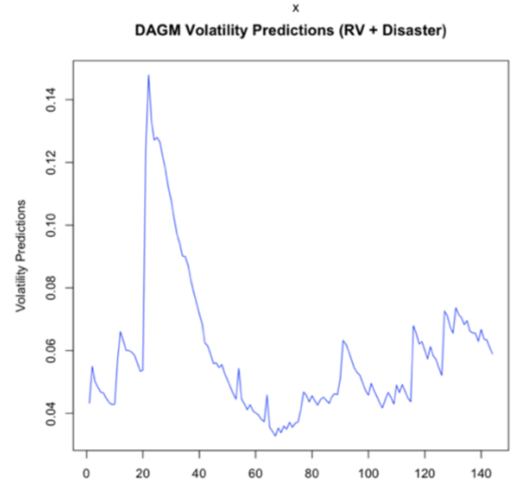

Volatility Predictions (RV)
Volatility Predictions
0.04 0.06 0.08 0.10 0.12 0.14
0 20 40 60 80 100 120 140
x


Volatility Predictions (RV + CPU)
Volatility Predictions
0.04 0.06 0.08 0.10 0.12 0.14
0 20 40 60 80 100 120 140
x


Volatility Predictions (RV + Disaster)
Volatility Predictions
0.04 0.06 0.08 0.10 0.12 0.14
0 20 40 60 80 100 120 140
X-axis Label


Volatility Predictions (CPU + Disaster)
Volatility Predictions
0.04 0.06 0.08 0.10 0.12 0.14
0 20 40 60 80 100 120 140
x


DAGM Volatility Predictions (RV + CPU)
Volatility Predictions
0.04 0.06 0.08 0.10 0.12 0.14
0 20 40 60 80 100 120 140
x


DAGM Volatility Predictions (RV + Disaster)
Volatility Predictions
0.04 0.06 0.08 0.10 0.12 0.14
0 20 40 60 80 100 120 140
x

Benchmark

For the out of sample forecasting, a rolling window approach was used – i,e, using data from January 1,1994 to December 31,2021 to project the price volatility on January 1,2022; using data from January 2,1994 to January 1,2022 to project the volatility on January 2,2022; and so on.

For the forecasting horizon, only 1 day forecasting horizon was chosen. Increasing the forecasting horizon decreased the forecasting abilities of the model as the model volatility dimishes over time.

To compare the predicted volatility models (GM and DAGM), a graph was plotted along with the actual realized volatility (Actual RV) from January 2022 – July 2022, as shown in the figure below.

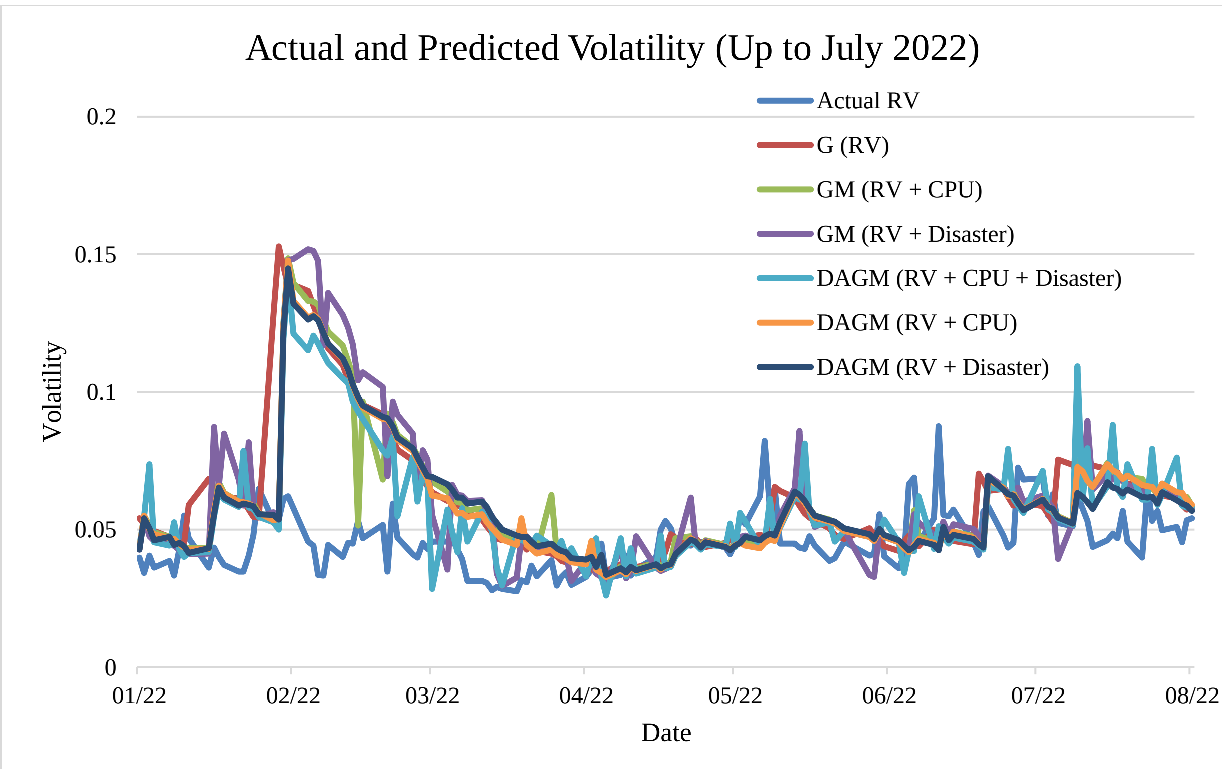


Figure 3.0 – Graph of Actual and Predicted Volatility

It is observed that all the predicted volatilities – whether from the GARCH, GARCH-MIDAS or Double Asymmetric GARCH-MIDAS models – generally predict higher volatilities than the actual volatility. This is most apparent during the period of February 2022, where the predicted volatility spiked to around 0.15 while the actual RV rose to around 0.06 only.

To calculate the Mean Squared Error (MSE) and Mean Absolute Error (MAE) of the predicted volatilities, we used 2 benchmarks, namely:

1. Actual RV
2. GARCH predicted RV, *G (RV)*

The computed errors of all the models are summarized in the tables below for both benchmarks:

Table 3a. Out-of-sample errors using Actual RV as benchmark.

Benchmark 1: Actual RV						
Method	**GARCH**	**GARCH-MIDAS (GM)**		**Double Asymmetric GM (DAGM)**		
Variable	**RV**	**RV + CPU**	**RV + Disaster**	**RV + CPU + Disaster**	**RV + CPU**	**RV + Disaster**
MSE	0.00085258	0.0007719	0.0009700	0.0006482	0.0007291	0.0007207
MAE	0.01953894	0.0185519	0.0203257	0.0177117	0.0183373	0.0180872

Table 3b. Out-of-sample errors using GARCH-predicted RV, *G (RV)* as benchmark.

Benchmark 2: GARCH RV					
Method	**GARCH-MIDAS (GM)**		**Double Asymmetric GM (DAGM)**		
Variable	**RV + CPU**	**RV + Disaster**	**RV + CPU + Disaster**	**RV + CPU**	**RV + Disaster**
MSE	0.0001665	0.0002099	0.0002224	0.0001465	0.0001506
MAE	0.0057907	0.0084036	0.0084865	0.0050666	0.0055955

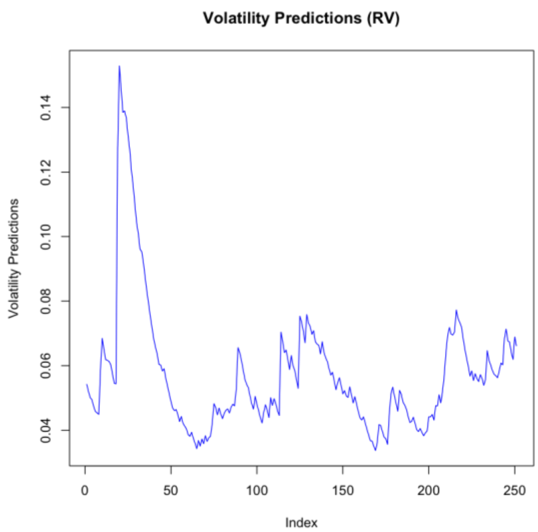
From observation, it is evident that the out-of-sample errors in Table 3b are less than that of Table 3a. This is also supported by Figure 3.0, where the predicted volatilities are quite close to each other and are generally higher than the actual RV. Generally, the Double Asymmetric GM can predict better than the regular GARCH-MIDAS model. In terms of comparing the predicted volatilities to the actual RV, the model that gives the least error (both MSE and MAE) is the Double Asymmetric GM using all three variables (RV + CPU + Disaster).

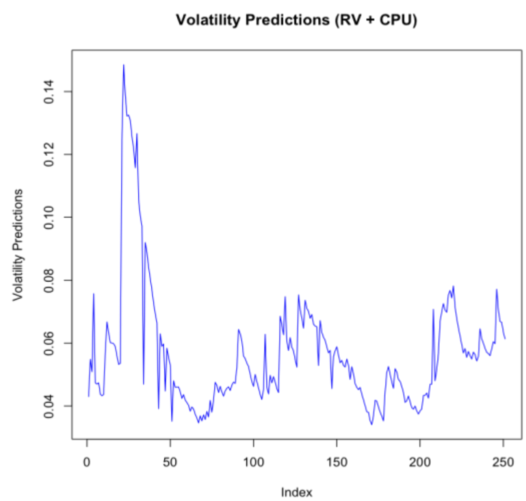
Extension

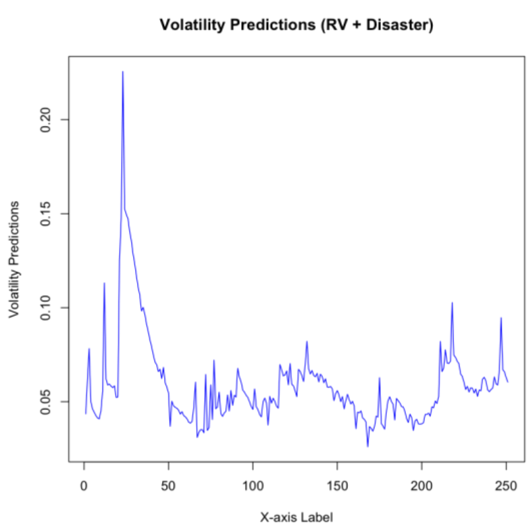
Extending our analysis from July 2022 to December 2022: We want to predict the volatility, which is driven by different combinations of long-term predictors such as return, CPU, and disaster. These predictors likely influence the volatility of the returns and provide insights into its performance and associated statistical characteristics. By extending the time interval and incorporating these predictors into the analysis, it would provide a more robust examination of the volatility prediction over time.

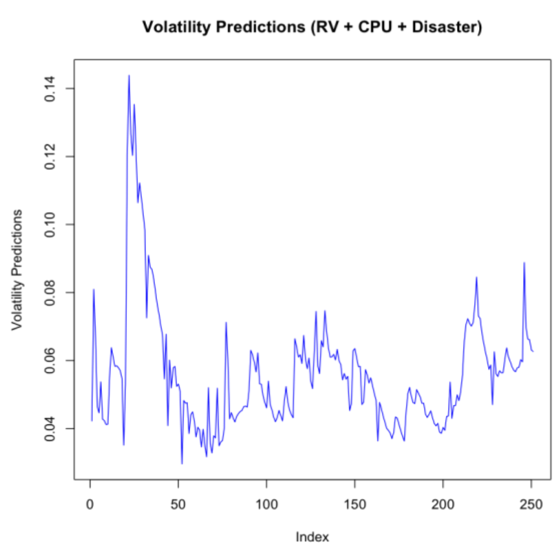
Graphical representation of following volatility predictions for the extended period:

- Return
- Return + CPU
- Return + Disaster
- Return + CPU + Disaster









We calculate the Mean Squared Error (MSE) to evaluate the predictive accuracy of our models when forecasting the future's returns within the extended timeframe.

Additionally, we also test the models using AIC, BIC, and QLIKE to further assess their adequacy and fit. AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) are model selection criteria that balance model goodness-of-fit with complexity, with lower values indicating better fit while penalizing for model complexity. On the other hand, QLIKE (Quasi-Likelihood) measures the quality of fit of a model based on quasi-likelihood estimation, with higher values indicating better fit to the data.

These metrics provide quantitative measures of the disparity between the predicted and actual values, offering valuable insights into the performance of our forecasting models and their ability to capture the underlying dynamics of the future.

Model results

MODEL	MSE	AIC	BIC	QLIKE	LLF
RV					14230
RV + CPU	0.0016633	-28732	-28684	-5.868857	14372.93
RV + DISASTER	0.0012750	-28486	-28438	-5.895262	14249.87
RV + CPU + DISASTER	0.0012725	-28481	-28406	-5.895774	14251.69

In comparing the models, **Model 4 (RV + CPU + DISASTER)** appears to be the best choice as it exhibits the lowest AIC and BIC absolute values.

Only the log-likelihood (LLF) favors Model 2 (RV + CPU) in this analysis.

While the MSE values are quite similar between Models 3 and 4, Model 3 shows a slightly higher absolute QLIKE, which can be interpreted as a better fit. Therefore, based on these criteria, Model 4 seems to be the optimal selection.

This confirms the significance of CPU and disaster frequency in predicting natural gas futures' price volatility in the extended period.

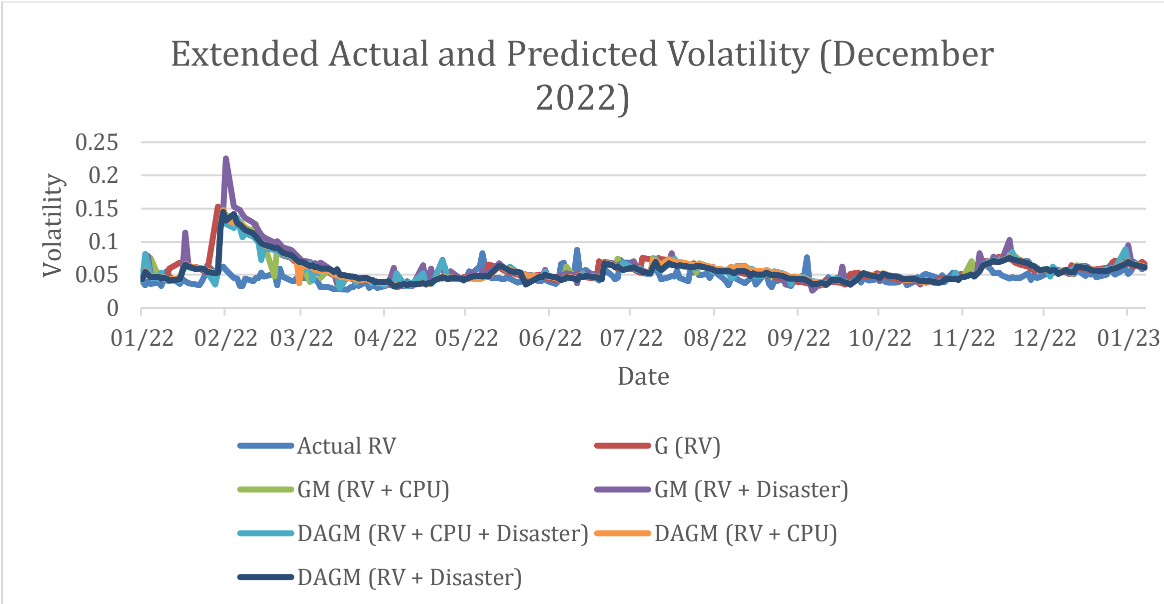


Figure 4. Graph of Extended Actual and Predicted Volatility

Table 4a. Extended Out-of-sample errors using Actual RV as benchmark

Benchmark 1: Actual RV						
Method	**GARCH**	**GARCH-MIDAS (GM)**		**Double Asymmetric GM (DAGM)**		
Variable	**RV**	**RV + CPU**	**RV + Disaster**	**RV + CPU + Disaster**	**RV + CPU**	**RV + Disaster**
MSE	0.00053978	0.00051906	0.00076104	0.00044860	0.00047831	0.00047681
MAE	0.01476771	0.01478487	0.01663133	0.01429373	0.01451331	0.01415596

Table 4b. Extended Out-of-sample errors using GARCH-predicted RV, *G (RV)* as benchmark

Benchmark 2: GARCH RV					
Method	**GARCH-MIDAS (GM)**		**Double Asymmetric GM (DAGM)**		
Variable	**RV + CPU**	**RV + Disaster**	**RV + CPU + Disaster**	**RV + CPU**	**RV + Disaster**
MSE	0.0001135	0.0001767	0.0001353	0.0000984	0.0000977
MAE	0.0053060	0.0072638	0.0062564	0.0047757	0.0047080

Similar to the results obtained during replication, the extended results also showed that the Double Asymmetric GM generally gives better volatility predictions than the regular GARCH-MIDAS model. In terms of MSE, the model that gives the least extended out-of-sample error is the Double Asymmetric GM using all three variables (RV + CPU + Disaster), which is consistent with the results obtained in the replication part.

Conclusion

The analysis in replication period reveals a discernible pattern in forecasting accuracy. Specifically, the RV+DISASTER model initially displayed the highest prediction errors among the suite of models considered. However, the incorporation of double asymmetry into this model markedly reduced these errors, highlighting the efficacy of this methodological refinement.

In a similar vein, the application of a standard GARCH model to RV data yielded the second highest level of forecasting errors. This outcome signals the limited predictive capacity of the GARCH model when used in isolation. Conversely, the integration of CPU and DISASTER variables into the forecasting framework considerably enhanced the model's predictive performance. Notably, the tri-variate RV+CPU+DISASTER model achieved the lowest error margins, establishing its superiority over other model configurations.

Crucially, these observed improvements were not confined to only the replication period but were also observed in the extended period. This enduring pattern underscores the utility of both CPU and DISASTER variables in bolstering the predictive accuracy of natural gas volatility models. The enduring nature of these results across different time frames solidifies the argument for including these variables in volatility forecasting models.

Acknowledgement

We would like to thank Prof. Lim Kian Guan for his continued support and guidance.

References

Predicting natural gas futures' volatility using climate risks

Rugarch Package (CRAN - Package rugarch (r-project.org))

Rumidas Package (CRAN - Package rumidas (r-project.org))

Adds the Mixing-Data Sampling (MIDAS, Ghysels et al.(2007) <doi:10.1080/07474930600972467>)

Components to a variety of GARCH and MEM (Engle (2002) <doi:10.1002/jae.683>

Engle and Gallo (2006) <doi:10.1016/j.jeconom.2005.01.018>

Amendola et al. (2024) <doi:10.1016/j.seps.2023.101764>)