QF603 Group 8 Project Report
16 Nov 2024

Andre LIM
Dylan LOO
Edmund CHIA
Eko WIDIANTO
Jacob LEE Zong Kang
Joseph ADHIKA

0. Thesis Statement

For our research project, we decided to explore Topological Data Analysis (benchmarked against other more traditional methods) to see if it yielded predictive power in forecasting market downturn and see what hedging strategies we could use to mitigate against market downside risk.

1. Background and Exposition

Topological Data Analysis (TDA) is a relatively modern tool in data science that analyses the shape (topology) of data. It can help uncover patterns and structural insights in financial time series that traditional methods might miss.

2. Literature Review

We would like to acknowledge a source material as literature review, reference for TDA and its applications for predicting market volatility : Gidea, M., & Katz, Y. (2017). Topological Data Analysis of Financial Time Series: Landscapes of Crashes. arXiv:1703.04385. https://arxiv.org/abs/1703.04385

3. Benchmark Indicator : VIX

For our traditional benchmark, we made use of the VIX as an indicator for predicting market downturn. In spite of the weaknesses of VIX as a leading indicator for market downturn, it is still commonly used. There are three prominent drawbacks about the VIX:

1. Lagging Indicator: it reflects past market conditions and may not provide timely signals for imminent downturns
2. False Signals: spikes in the VIX can occur without a corresponding market downturn, resulting in misleading interpretations about future market performance
3. Short-term Focus: the VIX is derived from short-term options pricing, making it less effective in capturing longer-term trends and downturns

Thus, this naturally implies that any novel methodology for forecasting severe market downturn should at least serve as a more reliable predictor of market downturn than the VIX index.

3.1. Predicting VIX using ARIMA

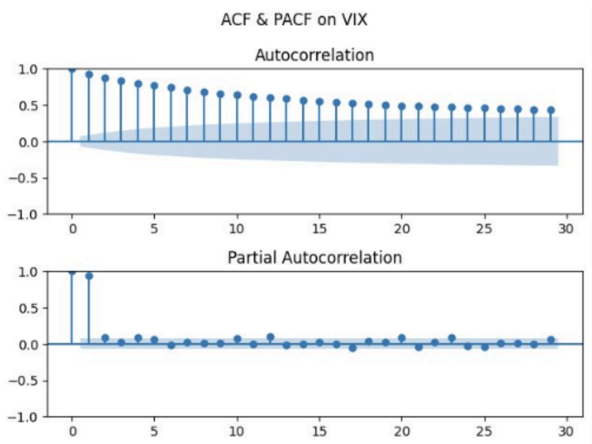
VIX data from 1994 until today are used in the modelling.

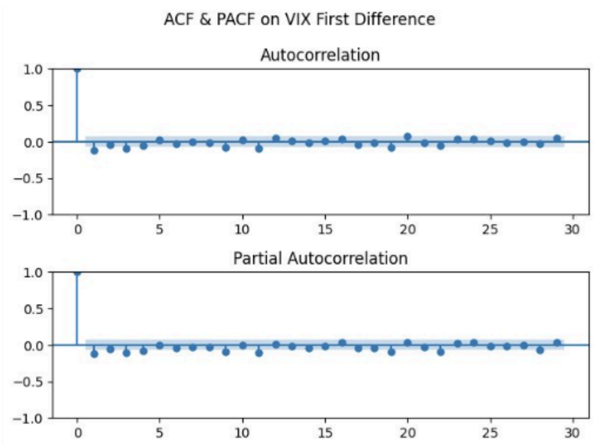
Autocorrelation Test	P - Value
KPSS	0.010
ADF	0.018

The KPSS test indicates non-stationarity while the ADF test indicates stationarity. It implies the series is difference-stationary. To achieve stationarity, the series should be differenced, and the resulting series should then be tested for stationarity.

p	d	q	AIC	LB (P-value)
1	0	0	1903.15	0.1810
0	1	0	1916.09	0.0480
0	0	1	2818.13	0.0000
1	1	0	1907.02	0.1858
0	1	1	1905.50	0.2286
1	0	1	1898.17	0.4178
1	**1**	**1**	**1886.09**	**0.9464**
1	1	2	1887.21	0.9531
2	1	1	1887.16	0.9529
2	1	2	1888.61	0.9541

Based on the lowest AIC value and high p-value (LB), ARIMA(1,1,1) is chosen.

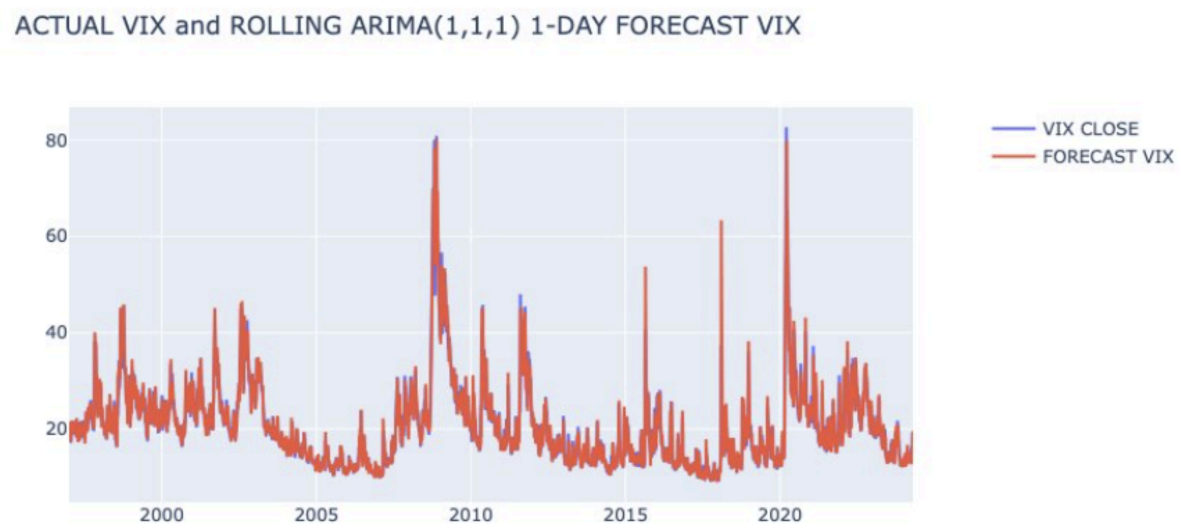




3.2 Rolling VIX Forecast and Hedging Event

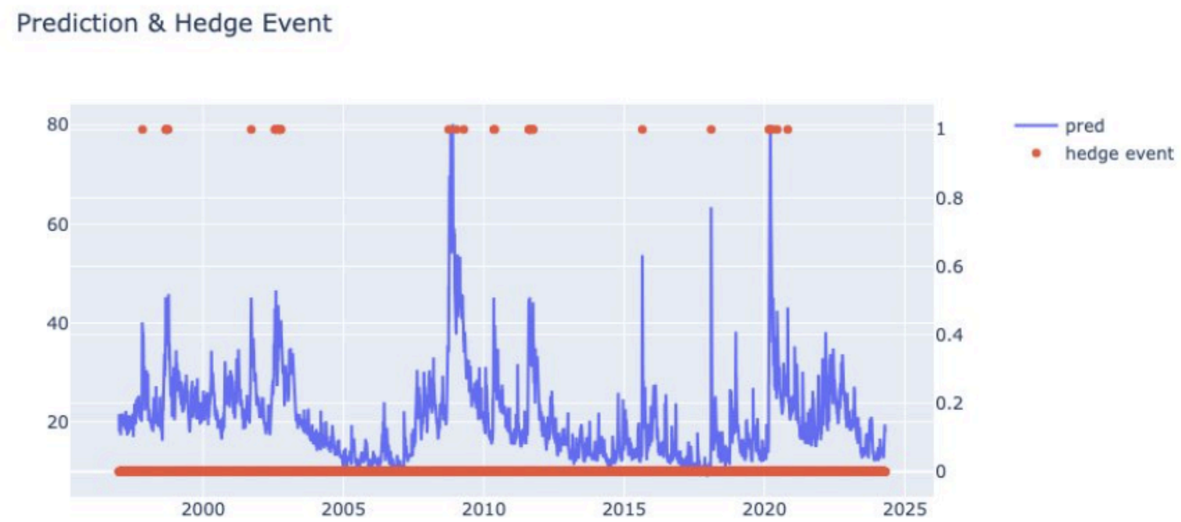
Rolling ARIMA Forecast:

- A rolling model is a predictive technique used in time series analysis that retrains the model at regular intervals and generates forecasts over a moving window of data.
- This approach adapts to changing data patterns, enhancing forecast accuracy and model robustness, especially for dynamic datasets.
- Continuous Retraining: The model is updated every set number of observations (20 days) to keep predictions relevant and responsive to new data.



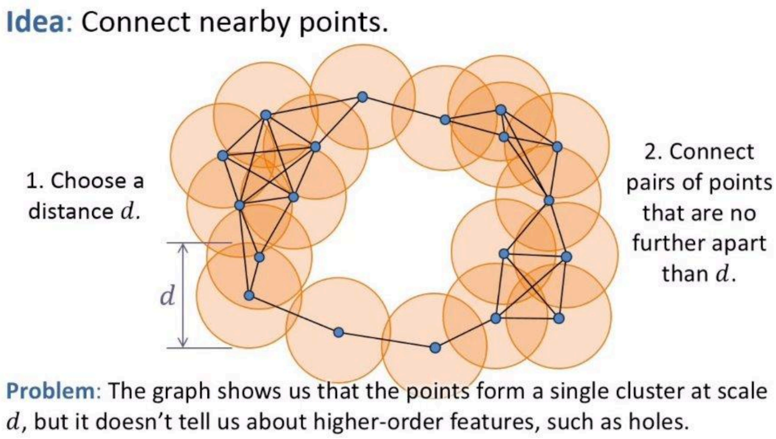
Hedging Event:

- Hedging signals are generated once the predicted VIX value is above its rolling long-run mean + 5 s.d.

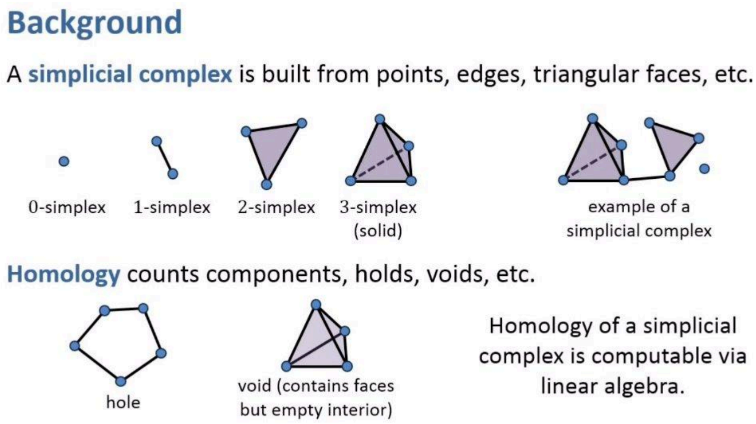


4. Persistent Homology

Persistent Homology is an algebraic method for discerning topological features of data. In the context of this project, data is a set of discrete points with a specific metric to measure distance between points and topological features of such data points may take on the form of components, holes and graph structures. For illustration diagrams pertaining to Persistent Homology, we would like to give due credit and acknowledgement to Matthew Wright.



When we imagine a set of point clouds in a 2D space, we need to choose a distance d such that if the parameters of point clouds of radius d intersect, it would co-join into a cluster, similar to a DBSCAN clustering algorithm.



A terminology we will often use in this report is the simplicial complex. A simplex can be thought of as a n-dimensional polygon. 0-simplex being a vertex, 1-simplex being a line, 2-simplex being a triangle, 3-simplex being a tetrahedron and so on. For the purpose of our project, we will be focusing on 2-simplex computations.

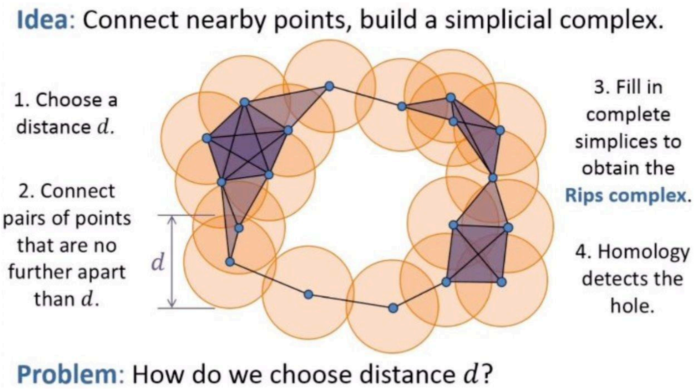
Homology is interested in the counting of non-trivial topological data structures like for example a hole (or a donut) given the distribution of data points. For the purpose of our exposition, we will be focusing on holes. Homology of a simplicial complex is computable via linear algebra.

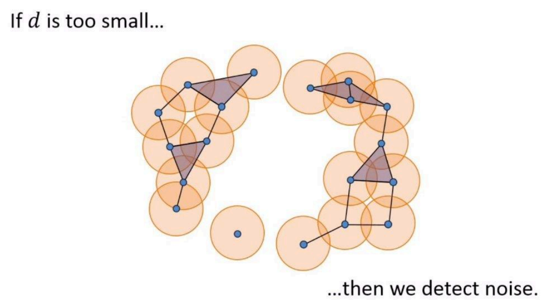
4.1. Rips Complex

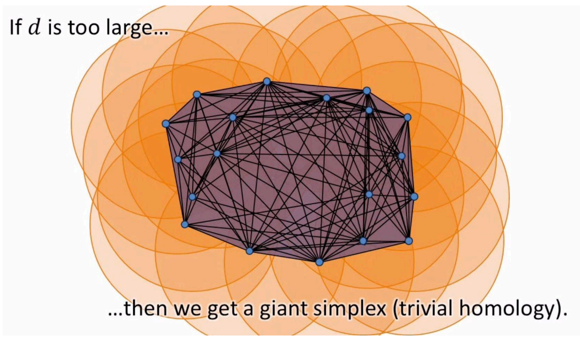
The Rips complex (also known as the Vietoris-Rips complex) is a type of simplicial complex that is particularly useful in Topological Data Analysis (TDA) for Persistent Homology. It is constructed from a set of points (a point cloud) by connecting points that are within a certain distance of each other.

For a given distance d, a Rips complex includes a simplex for every finite set of points that are pairwise within d of each other. So, if every pair of points in a set of three points is within d distance, these points form a 2-simplex (triangle) in the Rips complex.

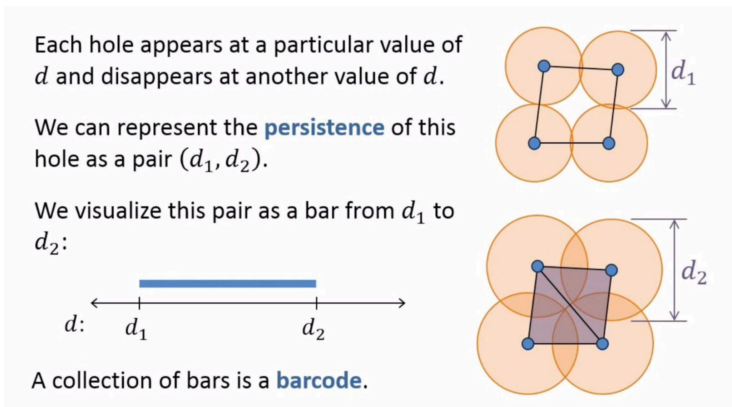
To achieve non-trivial homological features of data, we would need to compute optimal value d. If the d value is too small, we would retrieve many noisy outputs of mini-clusters. Conversely, if the d value is too large, we would retrieve 1 large giant simplex which is also trivial homology.



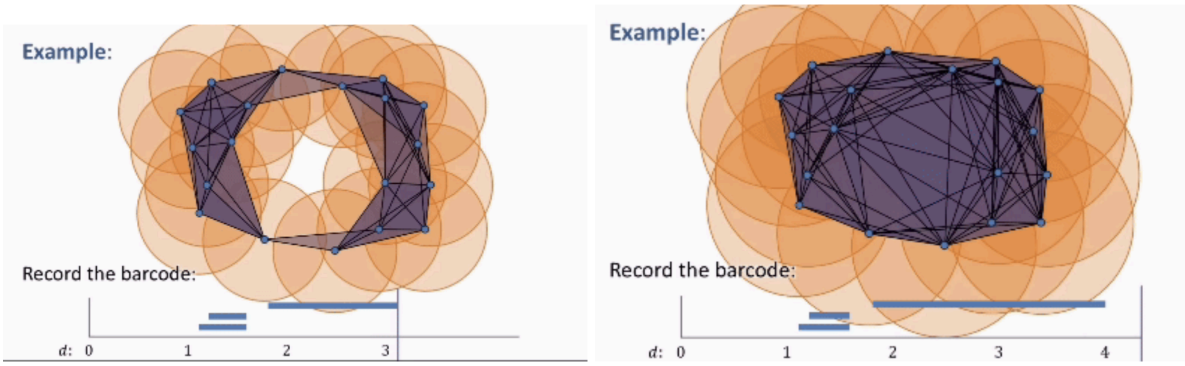




4.2. Barcode and Persistent Diagrams



We say that a homology diagram is persistent when its homology features endures over distance interval (d1, d2).

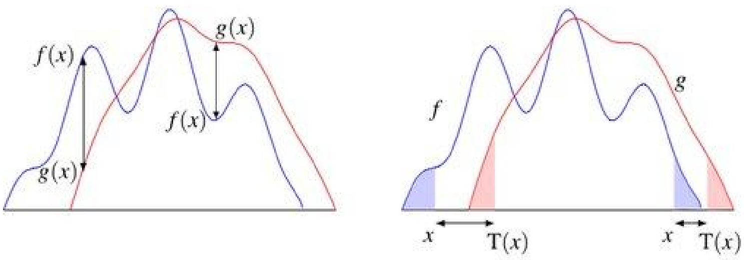


In the illustration above, the “donut” homology persists between d=2 and d=4.

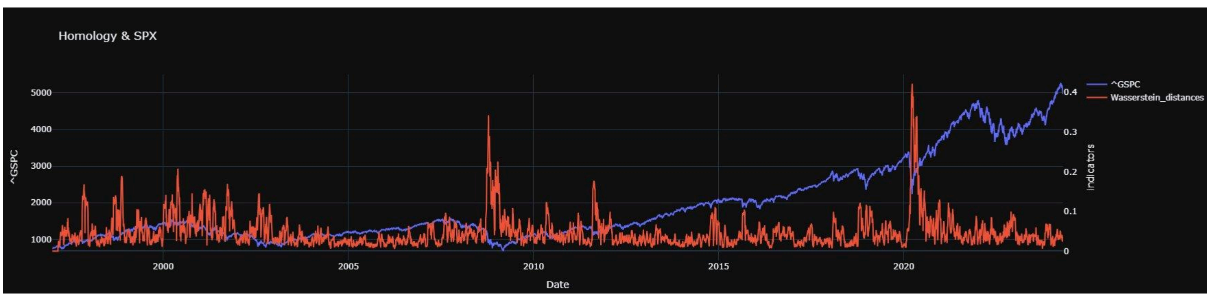
4.3. Wasserstein Distance

Wasserstein Distance provides a meaningful way to quantify the difference between two distributions. In simple terms, we can think of it by thinking of it as the cost of transforming one distribution into another through optimal transportation of mass (Earth mover's distance)

Sensitivity to the Structure of Data: Unlike other distances like the Euclidean distance, the Wasserstein distance is more effective as it considers the underlying geometry of the probability space. It reflects differences between distributions in a way that is more intuitive and aligned with human perceptions in many cases.



Source: Optimal Transport for Applied Mathematicians (F. Santambrogio)



4.3.1 Mathematical Intuition for Wasserstein Distance

So in our project, our program automatically computed persistence diagrams and summed the differences in Wasserstein Distance between said diagrams. Computation of persistent diagrams and optimization of d parameters for optimal non-trivial homology was also automated using the ripser.Rips python package. The intuition is that given a non-trivial homology (for e.g. a donut), as the homology evolves (i.e. donut changes shape), we wish to track the minimum sum of distance it would take to transform from one homology point cloud distribution to another (i.e. as the donut twists and changes, what is the minimum edit distance to reach from donut at T to donut at T+1).

4.4. Economic Intuition for using Wasserstein Distance

Change in Market Structure: A high Wasserstein distance indicates a substantial difference between the current topological structure of the market data and its historical norm. This could signify that the market is undergoing significant structural changes, which could be due to various economic factors such as shifts in investor sentiment, changes in economic policies, or reactions to global events.

Volatility and Market Turns: Persistent homology captures features like cycles and clusters within the data. A significant increase in the Wasserstein distance might reflect the emergence or disappearance of such features, suggesting increased market volatility or a potential turning point in market trends. For instance, the formation of new loops in the persistence diagram might indicate a cyclical behaviour that was not present before, potentially signalling a forthcoming downturn.

4.5. Other Computational Variables Complementary to Wasserstein Distance

Permutation Entropy

- Evaluates the **uncertainty** and **randomness** in the series by examining the order relations between values of the time series, rather than the values themselves.
- **Lower permutation** entropy might indicate a more **predictable** or **trending** market, whereas **higher entropy** could signal a **chaotic** or **volatile** market.
- Can be used to detect changes in market dynamics that precede major market moves, including downturns.

Hawkes Process

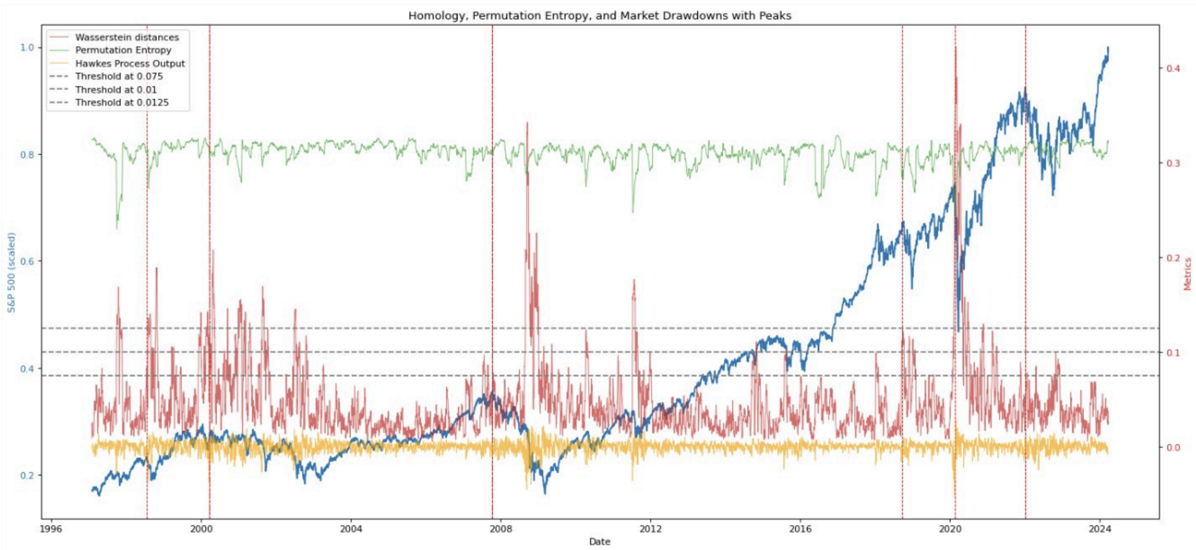
- The Hawkes Process is a type of statistical model known as a "self-exciting point process".
- It is characterized by the property that each event increases the probability of future events occurring in a short window of time.
- This is called **self-excitation** and it leads to clustering of events..
- Related to concepts like volatility clustering and the detection of catalyst activity.

4.5 Results (Wasserstein Distance)

For our project, our analysis focused mainly on the S&P500 daily data between 1997 to the end of Q1 2024. During this period we identified 6 significant drawdown periods:

1. 17 July 1998 (Asian financial crisis)
2. 24 March 2000 (Dot-com bubble)
3. 09 October 2007 (Global Financial Crisis)
4. 20 September 2018 (Sino-American trade war)
5. 19 February 2020 (Covid crash)
6. 03 January 2022 (Bear market)

Prediction Efficacy : 1997 to 2024 Q1e



We then do a linear sweep of different levels of Wasserstein Distances (WD) to determine trade-offs between precision and recall rates.

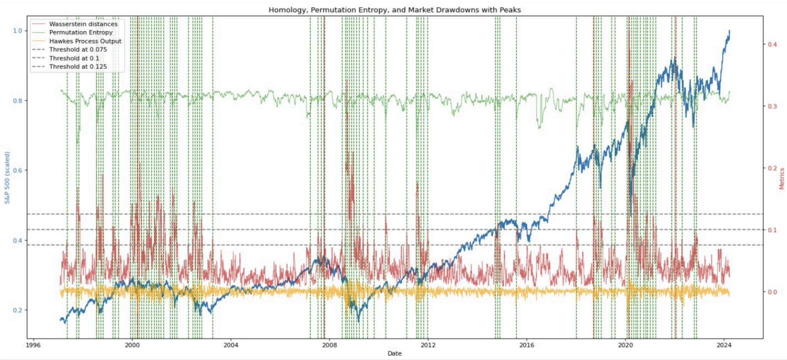
Precision: out of all prediction warnings, how many of them actually captured market downturn? ($0 \leq x \leq 1$, higher is better)
Recall: How many market downturns were captured out of all that occurred? ($0 \leq x \leq 1$, higher is better)
Simple DD (drawdown) Accuracy: measures proportion when value of time-series (TS) after 30-days is less than the value of TS when warning was alerted compared to the total number of times the alarm was sounded. Meaning to say that for the warning to be useful, the market had to dip within 30 days after receiving warning.

Wasserstein Distance 0.075 : Precision 0.107, Recall 1.0
Wasserstein Distance 0.1 : Precision 0.129, Recall 0.67 (missed 3 6)
Wasserstein Distance 0.125 : Precision 0.158, Recall 0.5 (missed 2 3 6)

Some challenges faced when trying to interpret the WD data results:

1. How to determine appropriate Wasserstein Distance threshold
2. How to interpret if leading or lagging indicator
3. How to interpret if warning signal is for bear market or flash crash
4. How to quantify market downturn

Prediction vs Actual Downturn (WD : 0.075)

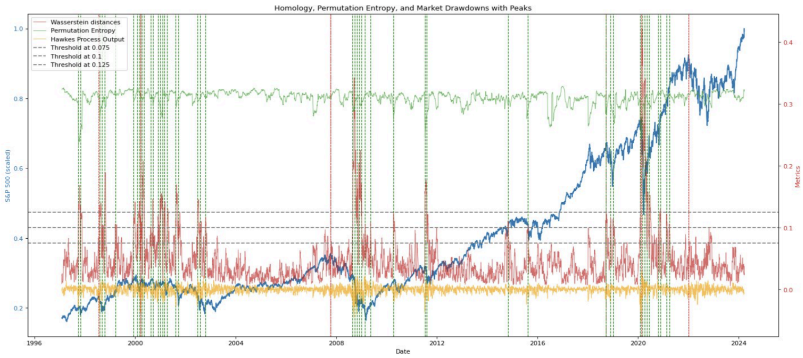


Actual Downturns (red vertical lines):

1. 1998-07-17
2. 2000-03-24
3. 2007-10-09
4. 2018-09-20
5. 2020-02-19
6. 2022-01-03

Precision: 0.107
Recall: 1.0
Simple DD Accuracy: 0.446

Prediction vs Actual Downturn (WD : 0.1)

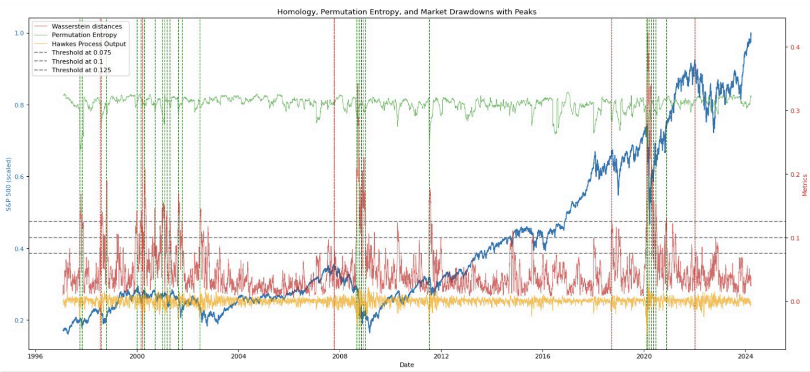


Actual Downturns (red vertical lines):

1. 1998-07-17
2. 2000-03-24
3. 2007-10-09
4. 2018-09-20
5. 2020-02-19
6. 2022-01-03

Precision: 0.129
Recall: 0.667
Simple DD Accuracy: 0.516

Prediction vs Actual Downturn (WD : 0.125)

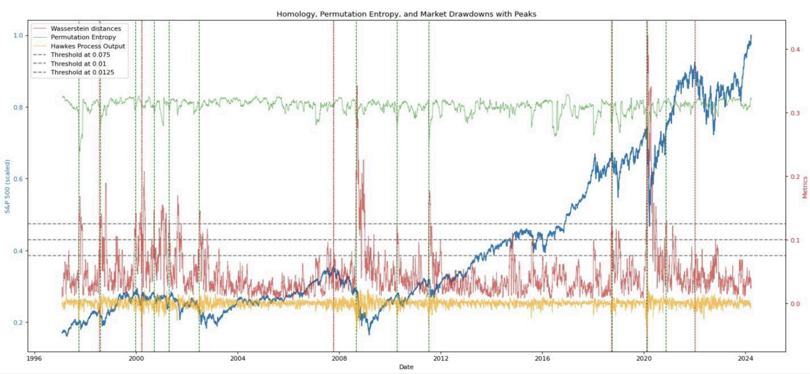


Actual Downturns (red vertical lines):

1. 1998-07-17
2. 2000-03-24
3. 2007-10-09
4. 2018-09-20
5. 2020-02-19
6. 2022-01-03

Precision: 0.158
Recall: 0.5
Simple DD Accuracy: 0.526

Prediction vs Actual Downturn (WD : 0.12 + cool-off 6M)



Actual Downturns (red vertical lines):

1. 1998-07-17
2. 2000-03-24
3. 2007-10-09
4. 2018-09-20
5. 2020-02-19
6. 2022-01-03

Precision: 0.25
Recall: 0.5
Simple DD Accuracy: 0.83

After doing a linear sweep, we eventually settled on WD of 0.12 with a cool-off period of 6 months (i.e. we ignore warnings for 6 months upon receiving the first warning)

5. Hedging : Methods and Data

5.1. Hedge Methodology:

The methodology involves selecting at-the-money (ATM) strike options with a minimum of 300 days to expiration (DTE). The strategies include holding the positions for 300 days and executing hedging when certain market conditions trigger (Hedge Flag = 1). Post-hedge analysis involves computing the portfolio returns and the 10-day 99% Value-at-Risk (VaR), which measures the potential loss in value of a risk-taking portfolio with a given probability, over a defined period.

5.2. Option Strategies:

The strategies discussed are designed to mitigate risks during significant market downturns:

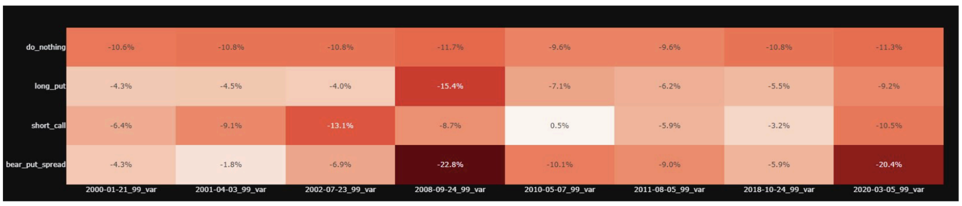
1. **Long Put**: This strategy involves buying put options to profit from or hedge against market declines. It serves as insurance against a downturn, as the increase in the value of put options during a decline can offset losses in the portfolio.
2. **Short Call**: This involves selling call options, which is typically beneficial in neutral or declining markets. The seller receives the premium from selling the options, which can be profitable as long as the market does not rise above the strike price of the calls.
3. **Bear Put Spread**: This is a more nuanced strategy that involves buying a put option at a higher strike price and selling another put option at a lower strike price. It limits both the potential loss and gain but is less costly than buying a single put due to the premium received from selling the lower-strike put.

5.3. Outcome Measures: 10-day 99% VaR

The effectiveness of each hedge strategy is quantitatively assessed using the 10-day 99% Value-at-Risk (VaR). This risk management tool measures the maximum potential loss in a portfolio over a specified period (10 days in this case) at a given confidence interval (99%). By calculating the VaR for each strategy during various historical periods, including market crashes, we now have a comparative analysis of how much protection each strategy offered against downside risk. The strategies were compared through statistical tests (e.g., t-tests) to determine if hedging significantly outperforms doing nothing.

Homology 10 Day 99%VaR

Total Hedges = 8

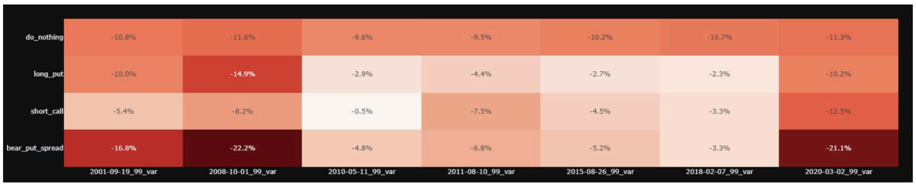


T-Test Difference of Means:
H0: Do nothing = Hedge
H1: Hedge > Do nothing

Strategy	P-Val
Long Put	1.56%
Short Call	2.53%
Bear Put Spread	42.9%

VIX/ARIMA 10 Day 99% VaR

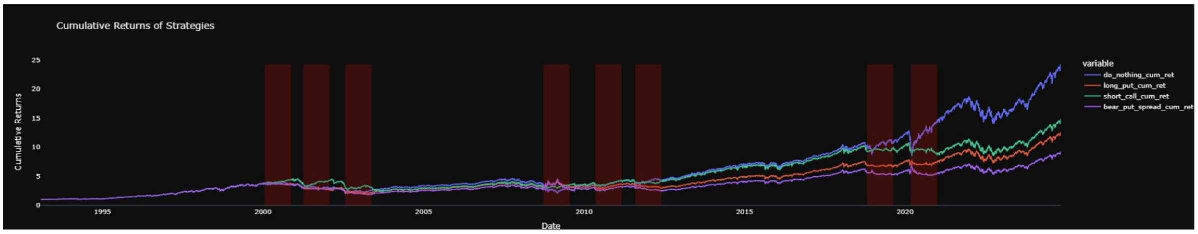
Total Hedges = 7



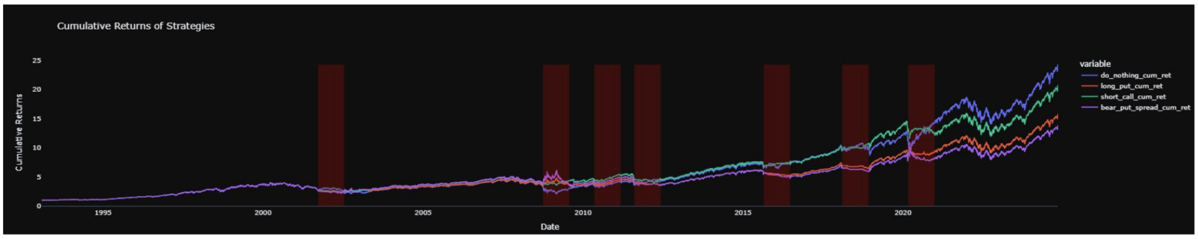
T-Test Difference of Means:
H0: Do nothing = Hedge
H1: Hedge > Do nothing

Strategy	P-Val
Long Put	4.63%
Short Call	0.99%
Bear Put Spread	61.36%

HOMOLOGY



VIX / ARIMA



5.4. Conclusion

The findings suggest that certain hedging strategies making use of persistent homology signal outperforms using VIX indicator when it comes to mitigating the risks during severe market downturns. However we also observed in both hedging scenarios, we lost out on the market rebound when our hedging period was too long. A possible area of further examination or improvement could be to check for market trend reversals and see if persistent homology may provide any insights into the same.