Stony Brook University

Applied Mathematics
&
Statistics

517 Report - Spring 2022

Modeling Maxima with an Autoregressive Conditional Frechet Model

Jason Bohne and Jarryd Scully

originally authored by Zifeng Zhao, Zhengjun Zhang, and Rong Chen

Contents

1 Introduction

Time Series Analysis is an ever-growing field of statistics with applications in biology, engineering, econometrics, and finance. One area in particular with a rich literature are series that can be characterized as Generalized Extreme Value Time Series. All G.E.V. series, Q_t, can be thought of as a function of an independent time series , X_t, with values in Q_t being determined by some exteremum present in X_t. Another way to interpret this as the application of Extreme Value Theory (E.V.T.) to traditional time series analysis.

The vast majority of literature so far focuses on static parameterizations of extreme values in time series. While the belief of time-independent parameters is fair in many fields, under financial applications this assumption is very strong. Two well known and documented phenomenon, volatility clustering and tail risk-clustering serve as empirical examples to why extreme value distributions should in fact be viewed as a function of time. That being said our project focuses on the academic paper **Modeling Maxima with an Autogressive Conditional Frechet Model** [1], originally authored by *Zifeng Zhao, Zhengjun Zhang, and Rong Chen.* Here the authors present a novel solution for modeling extreme value time series as a function of time. Specifically where the shape and scale parameter of the extreme value distribution (Frechet), follow an autoregressive equation of order 1. While this framework can be extended to other domains, careful consideration is given throughout the model description and results to provide financial examples and applications.

Our report proceeds as the following. In section 2, we will give a specification of the model. From there we will present the original simulations and data analysis the authors included, respectively in section 3 and 4. In section 5 we provide two of our own implementations of AcF models to both crypto orderbook data along with CBOE Volatility Index (VIX) data. Finally we share key takeaways in section 6, our conclusion. The appendix includes a detailed review of data collection, code access, and references.

2 AutoRegressive Conditional Frechet Model

2.1 Model Specification

We begin the specification of an AcF model with a given extreme value time series Q_t. Here we assume there exists some $(\mu_t, \sigma_t, \alpha_t)$ where $\alpha_t > 0, Y_t$ is i.i.d unit Frechet R.V. and $(\mu_t, \sigma_t, \alpha_t) \in \mathcal{F}_{t-1}$ s.t.

$$Q_t = \mu_t + \sigma_t Y_t^{1/\alpha_t}$$

In other words the value of the extreme value series at time t can be parameterized as a Frechet random variable of $(\mu_t, \sigma_t, \alpha_t)$. Moreover we assume μ_t is fixed across time s.t. $\mu_t = \mu$ and σ_t and α_t are of the form:

$$\log \sigma_t = \beta_0 + \beta_1 \log \sigma_{t-1} + \eta_1(Q_{t-1})$$
$$\log \alpha_t = \gamma_0 + \gamma_1 \log \alpha_{t-1} + \eta_2(Q_{t-1})$$

Where $\beta_1, \gamma_1 \geq 0$. By constructing an autoregressive equation of order 1, we can capture the desired time series effect in both the σ_t and α_t parameter values.

Note the conditional term in both processes. While the only conditions on $\eta_1(\cdot), \eta_2(\cdot)$ are to be continuous increasing functions, for simplicity let's use the following exponential:

$$a_0 exp(-a_1 x)$$

All together then our model is of the form:

$$Q_t = \mu + \sigma_t Y_t^{1/\alpha_t}$$
$$\log \sigma_t = \beta_0 + \beta_1 \log \sigma_{t-1} - \beta_2 exp(-\beta_3 Q_{t-1})$$
$$\log \alpha_t = \gamma_0 + \gamma_1 \log \alpha_{t-1} - \gamma_2 exp(-\gamma_3 Q_{t-1})$$

Where Y_t is i.i.d unit Frechet R.V.s , $0 \leq \beta_1 \neq \gamma_1 \leq 1$, and $\beta_2, \beta_3, \gamma_2, \gamma_3 > 0$.

2.2 Parameter Estimation

As we assume Q_t to follow a Frechet distriubtion of $(\mu_t, \sigma_t, \alpha_t)$, the conditional pdf can be derived. From there one can then compute the log-liklihood function as:

$$L_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} [\log \alpha_t + \alpha_t \log \sigma_t - (\alpha_t + 1) \log(Q_t - \mu) - \sigma_t^{\alpha_t} (Q_t - \mu)^{-\alpha_t}]$$

It can be shown that for any initial value (σ_1, α_1) there will always exist a unique, consistent, and asymptotically normal sequence θ_n that will maximize $L_n(\theta)$, i.e.

$$\theta_n = argmax_{\theta \in \Theta_n} L_n(\theta)$$

In other words, the MLE is well-defined and converges to best-fitted parameter estimates.

3 Original Results

3.1 Simulations

Before testing the AcF on historical data, several simulations were performed to analyse the performance of the model. The first study investigated the finite sample behavior of Q_t as shown in Proposition 1.

Proposition 1. *Given F_{t1}, denote $a_{pt} = 0$ and $b_{pt} = (\sum_{i=1}^{p} \sigma_{it}^{\alpha} t)^{1/\alpha_t}$*, we have, as p approach infinity,

$$\frac{Q_t - a_{pt}}{b_{pt}} \rightarrow \psi_{\alpha t}(x). \tag{1}$$

where $\psi_{\alpha t}(x)$ is a Frechet type random variable with tail index α_t and $\psi_{\alpha t}(x) = exp(-x_t^{\alpha})$

Using the limit, we can study the convergence of the marginal distribution of Q_t to the Frechet limit for the following one-time period factor model.

$$X_i = \beta_i \zeta + \sigma_i \epsilon_i \tag{2}$$

where Z follows a standard normal, β_i's are *i.i.d.* generated from a uniform distribution, and ϵ_i's are *i.i.d.* t-distributions with degrees of freedom . The σ_i's are generated from the uniform mixture distribution $0.5U(0.5, 1.5) + 0.5U(0.75, 1.25)$.

The original study used this model to compare the empirical distribution of Q and the limit stated in Proposition 1 for various pairs of vega and p. For each pair, 1000 sets of X_i's are generated, resulting in 1000 samples of the maxima process Q. The empirical CDF of Q was plotted to compare against the corresponding limiting Frechet distribution. Shown in Figure 1, the empirical distribution of Q approaches the limit.

The second application of the simulated data was used to investigate the AcF's ability to approximate the maxima process. A correlation coefficient between the true and estimated process

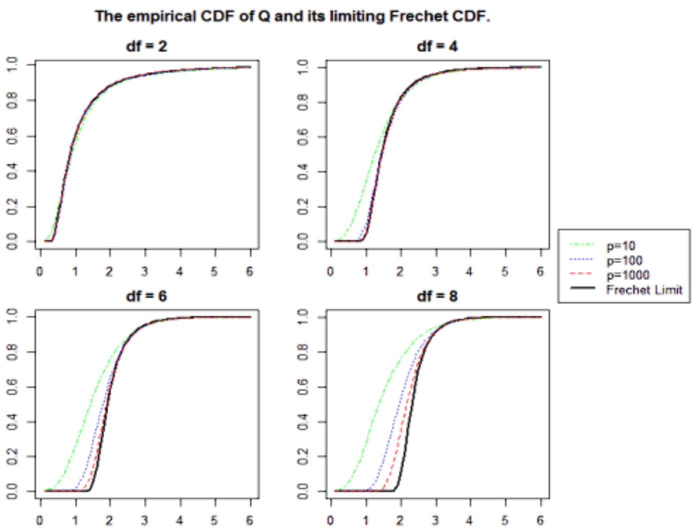

Figure 1: CDF of Q compared to the limiting Frechet CDF

was computed to quantify the accuracy of the approximation. In this section, the original study employed the following one-factor linear model.

$$X_{it} = 0.009(\beta_i \zeta_t + \sigma_i \epsilon_{it}), i = 1, ..., p; t = 1, ..., T, \tag{3}$$

where Z, β_i, and ϵ_i follow the previously stated distributions (see Equation 2). Differing from (2), ϵ follows a t-distribution with varying degrees of freedom $_t$ subject to the following series.

$$log_t = \gamma_0 +_1 log_{t1} + \gamma_2 exp(-\gamma_3 Q_t), \tag{4}$$

This specific series was chosen due to its ability mimic the tail index behavior demonstrated in AcF.

The AcF is used to model the simulated maxima process and the correlation between $_t$ and the estimated α is computed. The value for α was recovered using the fitted AcF. This coefficient represented the estimated tail index. The following correlations were computed for varying number of observations.

T_1	$\bar{q}$ $(q^0 = 0.1)$	$\bar{q}$ $(q^0 = 0.05)$	$\bar{q}$ $(q^0 = 0.01)$	mean cor.	median cor.
1000	0.095	0.049	0.012	0.871	0.928
2000	0.096	0.049	0.012	0.909	0.952
5000	0.097	0.051	0.012	0.947	0.973

Figure 2: Table of mean and median correlations.

The observed high correlation between the estimated and true coefficients indicates that the AcF is capable of detecting the evolution of the true tail index accurately.

3.2 Analysis of Major Stock Indices

After studying the capabilities of the AcF with simulated data, the investigation is continued using historical financial data. This section will analyze two sets of negative log return data for two major U.S. stock indices and later rate data from FOREX.

This section will analyze the maxima of the negative daily log-return data from the component stocks of the SP100 and DJI30 index. The return data was collected starting January 1, 2000 to December 31, 2014. For each trading day, the maxima negative log-return of each component stock was computed and represented by X_it. A daily maxima is computed among each of the components stocks and forms the maxima process Q_t. A maximum likelihood estimator (MLE) is performed to parameterize the fitted AcF model. The resulting coefficients are shown in the table below.

S&P100	γ_0	γ_1	γ_2	γ_3	β_0	β_1	β_2	β_3	μ
Mean	-0.068	0.890	0.328	5.33	-0.050	0.961	-0.051	6.68	-0.069
S.D.	0.014	0.013	0.063	1.27	0.006	0.004	0.0072	1.01	0.006
DJI30	γ_0	γ_1	γ_2	γ_3	β_0	β_1	β_2	β_3	μ
Mean	0.023	0.895	0.261	16.32	-0.052	0.964	-0.047	7.38	-0.059
S.D.	0.016	0.013	0.041	3.529	0.005	0.004	0.0066	0.813	0.006

Figure 3: Results for the model parameters.

The estimated parameter β_1 defines the σ_t process. The results shown a near unit value for β_1 which indicates a strong persistence of the σ_t series. The estimated tail index α_t was recovered from the model and plotted against the maxima process Q_t.

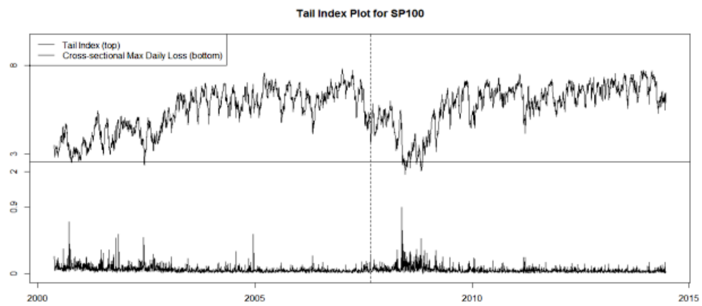

Figure 4: SP100 tail index plot.

Demonstrated in the figure above, the tail index is capable of following the general movement of the maxima process. The following MLE and tail index plots was repeated for the DJI30 maxima process. The result of this analysis are shown below.

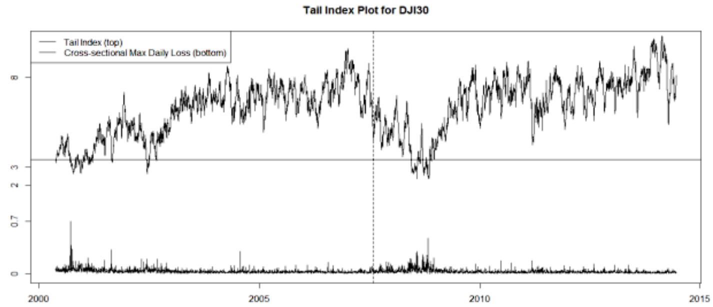

Figure 5: DJI30 tail index plot.

As discussed above, the estimated tail index for the AcF is capable of moving with the maxima process. The graph indicates a clear negative correlation between Q_t and the estimated tail index. The results from the two indices investigation shown that the AcF is capable of properly modeling the tail behavior of major stocks in the financial industry.

3.3 Analysis of Historical USD/JPY Exchange Rates

Similar to the last analysis, this section attempts to further analyze the performance of the AcF under historical finance data. This section utilizes 3-minute negative log-returns of foreign exchange data between the United States and Japan from 2008 to 2013. The 3-minute log-returns represents the series X_{it}, whereas Q_t series represents the daily maxima of X_{it}. MLE was used to parameterize the AcF. The results of this process are shown in the following table.

	γ_0	γ_1	γ_2	γ_3	β_0	β_1	β_2	β_3	μ
Mean	0.448	0.587	0.658	20.84	-0.120	0.890	-0.195	6.59	-0.051
S.D.	0.144	0.123	0.203	4.52	0.016	0.012	0.024	0.955	0.010

Figure 6: Model parameter results.

Considering the results of γ_1, the FOREX series is inherently more volatile than the previously studied index data. This suggests that high-frequency trading of currency exchange rates will present higher risk when compared to trading index components. To assess the performance of the AcF, a 1-day CVaR prediction was computed and compared to the true daily maxima Q_t. Additionally, a static GEV is included to understand the difference in results from a static and time-varying model.

Clearly shown above, the time-varying approach from the AcF better represented the true

		AcF		static GEV	
$q^0(\%)$	*Expected Violation*	*Violation*	*p-value*	*Violation*	*p-value*
10	61.6	60	0.89	32	0.00
5.0	30.8	35	0.41	17	0.01
1.0	6.2	8	0.41	2	0.10
0.5	3.1	4	0.56	1	0.39
0.1	0.6	0	1.00	0	1.00

Figure 7: Results for the 1-day CVaR prediction.

process Q_t. These results stress the importance of using a time-varying approach when modeling financial data.

4 Our Results

4.1 Crypto Orderbook Application

Our first real application utilizes orderbook data of the perpetual Bitcoin future, BTCUSDP, one of the most liquid instruments on the popular cryptocurrency exchange FTX. The parent time series X_t is composed of 3 months of 1-minute midprice data for the BTCUSDP market. For a further description on data collection and preprocessing performed, see the appendix.

Once the parent time series is formed, we can generate an extreme value time series Q_t, by computing the 60-minute max negative return of X_t. There are roughly 2150 observations within Q_t, which as expected is $\frac{1}{60}th$ the length of X_t. In figure 8, the midprice series X_t, return series $\bar{X}_t$ and the extreme value time series Q_t can be observed.

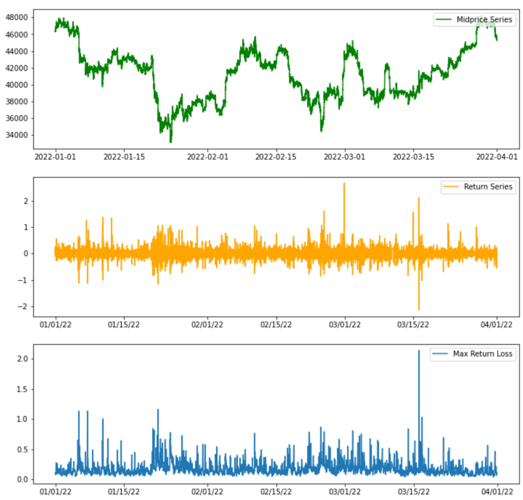

Figure 8: BTCUSDP Time Series

Now that we have the data prepared, we can create training and testing sets. Following in suit with the original work, let's fix a location parameter throughout our training set. To do this, we utilize Scipy's conditional MLE solver [2] that can be applied to a variety of distributions including the Frechet (Inverse Weibull) distribution.

Once we have a fixed location parameter we can calculate optimal scale and shape parameters for the second half of our training set (observations 750-1500). Iterating over the second subset, we perform conditional MLE at each timestep t to estimate most likely values for σ_t and α_t. Results can be seen in the figure below. 9

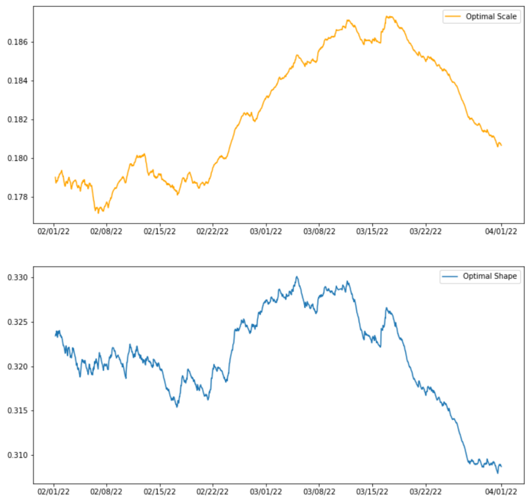

Figure 9: Optimal Time Varying Scale and Shape Parameters

Using the estimated σ_t and α_t values, we now calibrate the coefficient values for the autoregressive processes. Values are solved via the MLE solver provided in the Arch library [3]. Numerical results can be found below

Coefficients	**Numerical Estimates**	**[0.025, 0.975]**
β_0	0.0017	[-0.000 , 0.004]
β_1	0.9996	[0.998 , 1.001]
$\beta_2 * \beta_3$	-0.0028	[-0.003 , -0.003]
γ_0	-0.0010	[-0.004 , 0.002]
γ_1	0.9948	[0.992 , 0.998]
$\gamma_2 * \gamma_3$	0.0058	[0.005 , 0.006]
μ	-0.0529	N/A

Let's finally apply the model to the out of sample test set. Stepping through the remaining 650 elements we compute σ_t and α_t for each t based on the calibrated processes. Similar to the paper we then perform two goodness of fit comparisons. Specifically we compare the α_t process to the empirical CVAR [4] and the σ_t process to the average volatility under a GARCH(1,1) model. Results for the shape goodness of fit test are shown in figure 10. Note near unit correlations support the argument that the shape process is a descriptive measure of time-varying tail-risk.

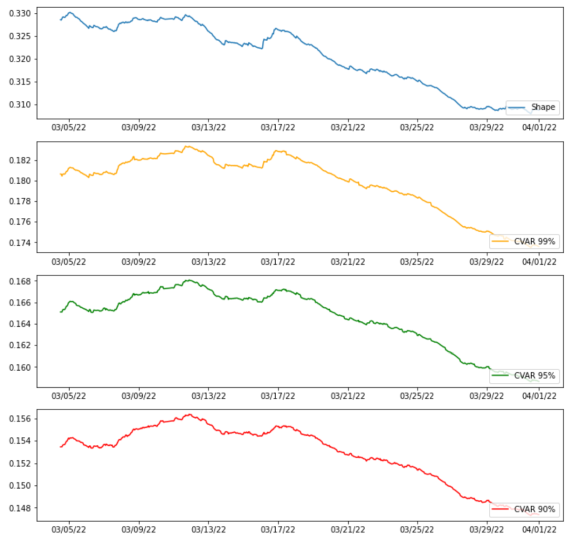

Figure 10: $Corr(0.1) = 0.9413$, $Corr(0.05) = 0.9407$, $Corr(0.01) = 0.9376$

4.2 CBOE Volatility Index Application

Our next investigation analyzed the AcF's ability to represent return data from the CBOE Volatility Index (VIX). The VIX is a commonly quoted index that represents the market's expectations for the relative strength of price changes in the Standards and Poor 500 index. A time series is formed from computing the difference between the daily closing price data. The series Q_t is formed by taking weekly maximums of the daily percent change in price. To gather enough data, a time frame of 1992 to 2022 was utilized.

Similar to the process described in the analysis of cryptocurrency data, Q_t is separated into a training and test set. Using the training set, we fit a Frechet Distribution using an MLE. The results are shown in the following table.

Parameters	Numerical Estimates
μ	-5.44
σ	5.489
α	178.33

Figure 11: Coefficient Results for the Training Set

Using the previously computed location parameter, the second half of the training set was used to fit a time varying shape and scale parameter. The results of this process is shown in the plot below.

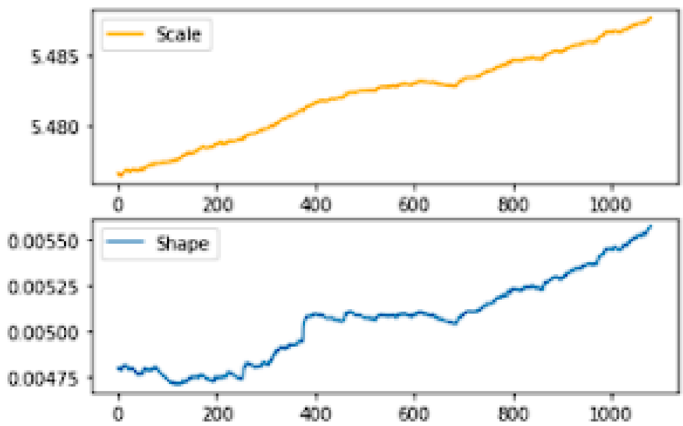

Figure 12: Plot of the Time-Varying Shape and Scale Parameters

For clarification, the plot for the shape parameter is one over the resulting coefficient. The following plot shows how the shape parameter will increase of time. This suggests that the tails of returns will increasingly get fatter over time.

Using the estimated σ_t and α_t values, the coefficient values for the auto-regressive process are estimated. As stated above, this process utilizes a maximum likelihood estimator.

Coefficients	**Numerical Estimates**	**[0.025, 0.975]**
β_0	0.0038	[0.003 , 0.005]
β_1	0.9979	[0.997 , 0.999]
$\beta_2 * \beta_3$	-0.0002	[-0.000 , -0.000]
γ_0	0.0108	[-0.005 , 0.026]
γ_1	0.9981	[0.995 , 1.001]
$\gamma_2 * \gamma_3$	-0.0223	[-0.025 , -0.020]
μ	-5.44	N/A

With the coefficients computed, the last step is to step through the test set for the remaining values of σ_t and α_t. To test the goodness of fit, the α_t process was compared to the CVaR based on the emprical data of the maxima. In addition, a correlation between the estimated and ture process was computed. These results are shown below.

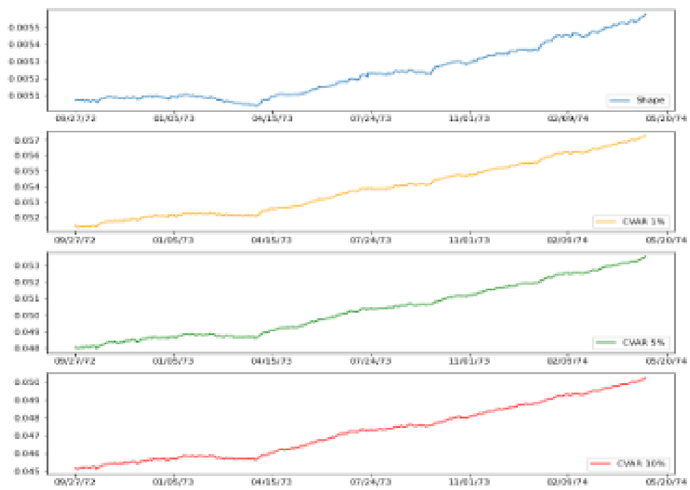

Figure 13: $Corr(0.1) = 0.9872$, $Corr(0.05) = 0.9875$, $Corr(0.01) = 0.9887$

The plot above clearly displays a positive correlation between the shape process and the empirical CVaR. A larger value of α_t will result in a distribution with fatter tails and a higher conditional value at risk. These results suggest that the AcF can properly model the returns and general tail behavior of the VIX.

5 Conclusion

Altogether, our work consists of an overview of the academic paper **Modeling Maxima with an Autogressive Conditional Frechet Model** along with original results. Beginning with a quick literature overview, we describe the motivation for a time-varying parameterization of an extreme value time series. From there we provide the model definition along with the MLE procedure to estimate model parameters. After the model is fully formulated, we present a series of simulations and data analysis the original authors completed. Finally we share our own results and their respective implications.

For both the crypto orderbook and CBOE VIX application, the goodness of fit test applied to the shape process was stronger than the test applied to the scale process. One reason for this is the ambiguity in how the authors computed the average volatility under a GARCH(1,1) model. Despite this, best estimates for the lagged scale and shape processes' coefficients were near unit values, supporting the known behavior of volatility and tail-risk clustering.

Practical applications of the AcF model in finance can be seen in a variety of areas. One broad area can be tail risk modeling, some topics of which could be VaR/CVaR estimation or tail correlations. In particular, practitioners would find the shape process, α_t, to be of interest. Another area of application is volatility modeling and threshold estimation, where the scale process, σ_t, could be utilized. With the widespread availability of computational libraries in python tailored to time series and MLE, interested readers can easily replicate the results of this report. Our scripts and datasets are open-sourced for testing and can be accessed from the Appendix.

6 Appendix

6.1 Data collection techniques

For the application utilizing crypto orderbook data, three months of 1-minute midprice data is used for the perpetual Bitcoin future BTCUSDP. The date range for the data spans January 1st, 2022 to April 1st, 2022. Note here the 1-minute midprice is computed by taking the mean of all midprice tick values within that minute. Moreover each midprice tick value is generated as the average of the first level of each orderbook snapshot as below.

$$Midprice_i = \frac{Bid_{i,0} + Ask_{i,0}}{2}$$

6.2 Code Access

The full repository of all scripts and data utilized in this project can be accessed from our **GitHub**. Note the csv files containing the raw data for both applications can be found **here**

References

[1] Z. Zhao, Z. Zhang, and R. Chen, "Modeling maxima with autoregressive conditional frechet model," *Journal of Econometrics*, 2018. [Online]. Available: https://doi.org/10.1016/j.jeconom.2018.07.004.

[2] "Scipy 1.0: Fundamental algorithms for scientific computing in python," *Nature Methods*. [Online]. Available: https://doi.org/10.48550/arXiv.1907.10121

[3] K. Sheppard, "Autoregressive conditional heteroskedasticity." [Online]. Available: https://github.com/bashtage/arch

[4] Quantopian, "Empyrical: Common financial risk metrics." [Online]. Available: https://github.com/quantopian/empyrical