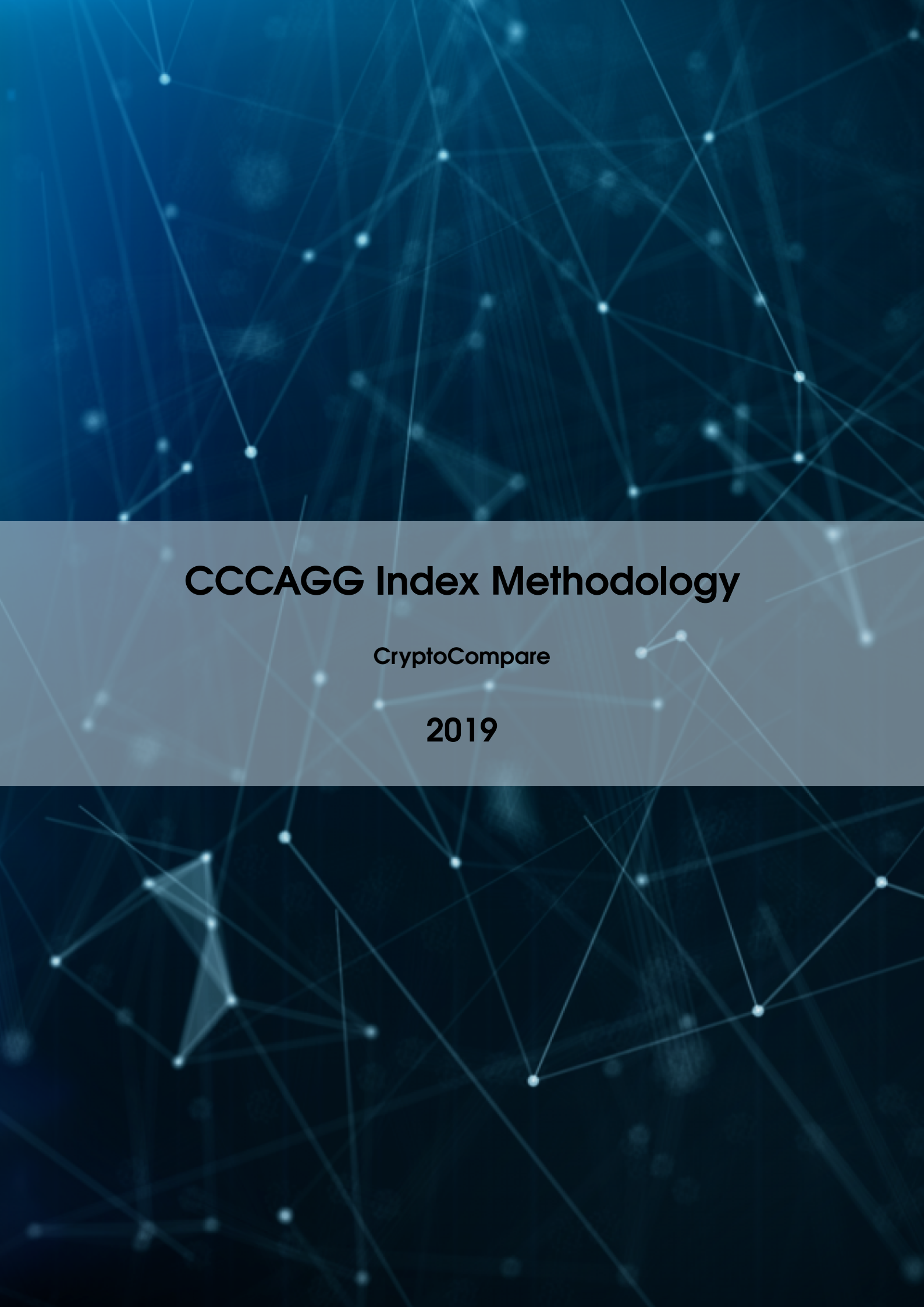

CCCAGG Index Methodology

CryptoCompare

2019

Contact info@cryptocompare.com

Published by CryptoCompare

CRYPTOCOMPARE.COM

17 October 2019

Contents

1. Version History

Version	Date	Reviewed by	Details
1	2017-07-01	Vlad Cealicu	Initial version
2	2017-11-20	Quynh Tran-Thanh	Added review process description
3	2018-02-26	Quynh Tran-Thanh, Konrad Strachan	Outlier methodology
4	2018-08-30	Ben Hunter-Craig	Amendment
5	2019-09-12	James Huckle	Outlier methodology update
6	2019-09-27	Ajay Pethani	Constituent exchanges selection methodology

2. Summary

2.1 Introduction

The Crypto Coin Comparison Aggregated Index ("CCCAGG") refers to the real-time index calculation methodology, the purpose of which is to show the best price estimation for crypto traders and investors to value their portfolio at any time. It aggregates transaction data of over 70 exchanges, using 24 hour volume weighted average. The CCCAGG is calculated for each crypto coin in each currency it is trading in (example: CCCAGG BTC-USD).

Crypto coins such as Bitcoin, Ethereum, Monero, etc. are traded at various markets against multiple currencies including fiat currencies (USD, JPY, GBP, etc) and other cryptos. Depending on the market type (exchange or OTC), liquidity level, trading volume, transaction fees, and many other factors, a coin can be traded at different prices across different markets, and therefore making it difficult to know the value of a coin at a certain time.

Unlike traditional stock exchanges, crypto exchanges are facing the following problems that make pricing even more challenging:

- DDOS attack, causing an interruption of trading
- Hacking of user accounts
- Lack of standards and naming convention for symbols
- Unstable technological and legal environment (causing changes in fee structure, blocking of funds withdrawal, etc).

2.2 Use and attribution

2.3 Description

Name	CCCAGG
Calculation agent	Crypto Coin Comparison Ltd.
Dissemination	Real-time
Day close	12 am GMT
Methodology	24 hour volume weighted average with time adjustment

3. Data collection

3.1 Data source

Transactional data (historical trades) is collected from each exchange via public REST API polled every 2 seconds or websocket endpoints. All collected data will be standardized internally, stored and backed up in servers.

Exchanges and markets are added on an ongoing basis based on research or user request. Exchanges that do not meet the technical requirements (available API for transactional data) cannot be added to the data collection. Unlike many data providers, who use snapshot data, CCCAGG approach of using transactional data enables auditability and replicability.

3.2 Data format

The collected data consist of:

- Trade id
- Timestamp (UNIX)
- Price
- Amount
- Position (buy/sell)

3.3 Failure of data retrieval

On the event of failure to retrieve data from an exchange (due to service outage on the exchange), per design of the CCCAGG indices, the last price of the respective exchange will expire over time (its weighting will decrease to close to zero). As long as the currency pair is trading on other exchanges, the CCCAGG calculation is uninterrupted.

4. Methodology

4.1 Input data

CCCAGG is calculated every time a new transaction gets executed. The following input data is needed from each transaction:

- Trade price
- Trade amount
- Trade timestamp
- Exchange where the transaction was executed

4.2 Constituent exchanges

Crypto Coin Comparison Ltd. has integrated with a list of exchanges, but only a subset of them count towards the calculation of CCCAGG.

The rule of thumb is to include as many exchanges as possible after a testing period. Exclusion happens in the following cases:

- Volatile prices compared to market average (OTC markets excluded)
- Exchange suspends trading activity
- Verified user or social media reports of fake data reporting
- Malfunctioning API (on the exchange side)

Each case for exclusion is discussed by the Review Committee on an ad-hoc basis but at least once every calendar month.An excluded exchange can be re-included if the Review Committee finds that the problem causing the market disturbance has been solved. Changes in constituent exchanges are communicated via API newsletter and direct communication channels with clients.

4.3 Index calculation

4.3.1 Volume weighted average

CCCAGG uses 24 hour volume weighted average to calculate prices. 24 hour volumes are calculated solely based on transactional data. This way CCCAGG is giving the most liquid market prices more importance, and price impact of illiquid (and therefore more volatile) markets are negligible.

4.3.2 Time adjustment

Next to volume weighting, a time penalty factor is also added, to ensure that exchanges that suspend trading has an expiring price impact. An example of a case where this methodology showed its advantage was the Bitfinex hack in 2016.

Bitfinex had one of the highest trading volumes in Bitcoin, therefore had a significant weight in most price indices. Therefore, when trading was suddenly suspended on Bitfinex, causing a crash on all other markets, most indices still showed the a Bitcoin price close to the last price on Bitfinex, although markets have already moved on.

CCCAGG takes last trade time into account, therefore the last Bitfinex price expired with time and the index could move with the market.

4.3.3 Aggregation over trading currency

CCCAGG only takes direct trading pairs into consideration for calculation. For example CCCAGG BTC-USD only accepts trades from exchanges trading BTC-USD directly, therefore no conversion needed for the aggregated index calculation.

The reason for this methodology is that a coin can trade on multiple currency markets with a significant price difference (premium or discount), therefore aggregating across all markets will result in an average price that is not useful for a trader or investor who holds a crypto position in a certain currency and will most likely trade in that currency.

4.3.4 Mathematical representation

For simplicity, we define the CCCAGG Price Index for a pre-speficied currency pair ABC_XYZ, all notations below refer to this pair.

Variable	Explanation
i, j	Exchange
t, s	Unix timestamp in seconds
P_t^i	Price at exchange i and time t
P_t^*	CCCAGG Index price at time t
W_t^i	Index weight of exchange i at time t
V_t^i	24 hour volume of exchange i at time t
x	Excluded from pricing, $x = \{0, 1\}$
γ_t^i	Liquidity penalty factor of exchange i at time t
h_t	Last hour timestamp at time t
τ_t^i	Time since last trade in minutes on exchange i at time t
δ_t^i	Price deviation of trade on exchange i at time t
A	Price deviation threshold

At any time t the Index price for a currency pair is the weighted average of prices on all exchanges.

$$P_t^* = \sum_i P_t^i W_t^i \tag{4.1}$$

A weight belonging to an exchange is the ratio of the 24 hour trading volume of a currency pair to the total volume of the pair on all exchanges.

$$W_t^i = \frac{V_t^i 1_{\{x=0\}} \gamma_t^i}{\sum_k V_t^k 1_{\{x=0\}} \gamma_t^k} \tag{4.2}$$

The 24 hour volume of an exchange at a certain time is defined as the sum of trading volume of the last 24 full hours.

$$V_t^i = \sum_{s=h_t-86400}^{h_t} V_s^i \tag{4.3}$$

The exchange volume is adjusted with the liquidity penalty factor, that decreases with the

increasing time since the last trade.

$$\gamma_t^j = \begin{cases} 1 & \text{if } \tau_t^i \leq 5 \\ 0.8 & \text{if } 5 < \tau_t^i \leq 10 \\ 0.6 & \text{if } 10 < \tau_t^i \leq 15 \\ 0.4 & \text{if } 15 < \tau_t^i \leq 20 \\ 0.2 & \text{if } 20 < \tau_t^i \leq 25 \\ 0.001 & \text{otherwise} \end{cases}$$

4.3.5 Outlier Detection

Aside from our qualitative based assessment on exchange suitability (discussed in 4.2), we also utilise both real-time and historical outlier detection mechanisms to exclude certain trades. This plays a crucial role in maintaining an accurate Index, because any outliers can have a sizable impact on the CCCAGG value.

For a trade to be considered an outlier, it must deviate significantly either from the median of the set of exchanges (historical), or from the previous index price (real-time). This kind of deviation could occur for a number of reasons - for example:

- Low liquidity on a particular instrument
- Erroneous data from an exchange
- Incorrect mapping of an instrument

The exact specifications of the outlier handling logic are defined below:

Outlier Detection: Real-time

The following outlier handling logic is implemented:

1. For one or two exchanges - no exclusion logic is applied. Everything is considered an inlier and included.
2. For three or more exchanges - Any trade that deviates by more than a certain threshold from the previously calculated index price (either up or down) causes the exchange to be labelled as an outlier and excluded.

The deviation is calculated for each new trade in the following way:

$$\delta_t^j = \begin{cases} \frac{1}{\frac{P_t^i}{P_{t-1}^*}} - 1 & \text{if } P_t^i < P_{t-1}^* \\ \frac{P_t^i}{P_{t-1}^*} - 1 & \text{if } P_t^i \geq P_{t-1}^* \end{cases}$$

If δ_t^j exceeds a threshold, a trade is deemed an outlier. The threshold, A is currently set as $A = 3$ For exchange j:

$$x_t^j = \begin{cases} 0 & \text{if } \delta_t^j < A \\ 1 & \text{if } \delta_t^j \geq A \end{cases}$$

This means that trades which exceed 400% or drop below 25% of the previous CCCAGG Index price will be deemed an outlier.

An edge case can exist where a split in the market occurs between two distinct groups of exchanges. If Group 1 exchanges trade at a significant price different to be included as outliers, only price and volume information from Group 2 will be included.

In this situation, if the number of exchanges in Group 1 (outliers) exceeds the number of exchanges in Group 2 (inliers), an inlier/outlier flip occurs with the previous included exchanges in Group 2 becoming excluded and the previously excluded Group 1 exchanges becoming included.

When an exchange is re-included - only current and future pricing data will reflect the re-inclusion. No back-filling or recalculation occurs for previously reported pricing in the Index.

Outlier Detection: Historical

1. For one exchange - no outlier logic is applied. Everything is considered an inlier and included.
2. For two exchanges - if the price of their respective last trades differs by more than a factor of two, the trade that is furthest away in percentage terms from the previous days' CCCAGG price - at midnight - will be considered an outlier and excluded from the calculation. If there is no price for the previous day, no outlier logic is applied.
3. For three or more exchanges - the median price of the set of exchanges' last trades is used as a benchmark to filter out any single last trade that differs from it by more than a factor of two.

$$M_t = median([P_t^i \text{ for exchanges } i])$$

$$x_t^j = \begin{cases} 1 & \text{if } 0.5M_t \leq P_t^j \leq 2M_t \\ 0 & \text{otherwise} \end{cases}$$

The following points are of note for all outlier handing:

(a) Exchanges with no 24 hour trade volume are excluded from all calculations.
(b) Excluded exchanges are exempt from the price weighting as well as the total volume.
(c) The exclusion of an exchange can impact the current price of the Index and consequently cause it to deviate from the previous Index price.
(d) An Exchanges' time adjustment and 24 hour volume will not be affected during a period of exclusion.
(e) Due to the outlier detection being calculated on every new trade arrival, excluded exchanges are automatically re-included in the Index if the Index average price moves to a point where the excluded exchanges' price is no longer considered an outlier.
(f) When an exchange is re-included - only current and future pricing data will reflect the re-inclusion. No back-filling or recalculation occurs for previously reported pricing in the Index.

4.4 Auditability and replicability

CCCAGG is auditable and replicable since its calculation is based on transaction data retrieved from exchanges via public API. Anyone who has access to this data can recreate the CCCAGG indices.

5. Review

5.1 Model review

The methodology is reviewed at least every quarter by the Review Committee. As of date the Review Committee consists of:

- Vlad Cealicu
- Quynh Tran-Thanh
- Constantine Tsavliris
- James Huckle
- Ajay Pethani

5.2 Exchange review

Exchange exclusion is reviewed by the Review Committee at least once every calendar month or on an ad-hoc basis, when market or technical events require. Market or technical events can be

- Suspended trading
- False data provision
- Service outage

In order to select constituent exchanges for a given pair, we use one month's worth of hour-granularity OHLCV data for all exchanges we have integrated with for the pair in question.

For each hour, we take the median of the hour-open price across the exchanges. For each exchange's price, we record the absolute percentage difference between the open of the hour, and that hour's median. We then take the average across the month, so for each exchange we have an average absolute percentage deviation. We also record the proportion of the total volume during the month accounted for by each exchange. Finally, we factor in the latest grade attained by each exchange using the methodology explained in our benchmarking report.

If a pair is only traded on at most 2 exchanges, then all exchanges for that pair are included in CCCAGG. If a pair is traded on more than 2 exchanges, then we use the above metrics as follows:

- All exchanges with grade E, F, or ungraded, are excluded, and AA to A grade exchanges are included automatically. If this results in fewer than 3 exchanges, we include grade E exchanges, and if this is still insufficient, we include F, and then ungraded.
- All exchanges with an average absolute percentage deviation exceeding 50% are excluded.
- All exchanges with a volume percentage of below 0.01% are excluded.
- After applying these filters, if we only have 0 or 1 exchange left, than we rank the exchanges by grade and then percentage volume, and add exchanges from the top of this list until we have 2 exchanges, and include only those two.

5.3 Change communication

If the Review Committee requires changes in the methodology or exchange exclusion, CCCAGG consumers will be notified via the API newsletter or will direct client communication channels.

6. Dissemination

CCCAGG indices are disseminated via REST API and Websocket API. Read the API documentation at https://min-api.cryptocompare.com/documentation.

7. Disclaimer

Crypto Coin Comparison Ltd is a registered company in England with a register number 08846965 and address 1 Duchess Street, W1W6AN, London, United Kingdom.