Automated Trading Model Research

Which source is the best at providing exchange data for the current automated trading model application?

Author: Nino Verhaegh

Group: Stockbrood

Date: 24-4-2023

Contents

Preface

For our project, we are making an application which tests the user's trading strategies by slightly changing it's variables. This eventually results in a lot of test result where the user can find the "sweet spot" for reach variable in order find the best strategy. an external software that will be used for this will most likely be MetaTrader 4. For the testing of these strategies, we want to run them through historical data of various FOREX exchanges. We now want to answer the research question: " Which source is the best at providing exchange data for the current automated trading model application?". Is this something that's built into MetaTrader 4? or are there api's that will give us this functionality?

1. What are the available sources for accessing exchange information?

To answer this question, the first DOT method I'll be using is **literature study**

Built in providers

MetaTrader 4 provides built-in functionality for backtesting trading strategies using historical data. It includes a strategy tester that allows you to test and optimize your trading strategies using a variety of parameters and indicators.

However, when it comes to obtaining exchange information, MetaTrader 4 relies on data feeds provided by third-party sources. You can choose from a variety of data providers that offer different levels of granularity and coverage.

Some popular data providers for MetaTrader 4 include:

- MetaQuotes Software: This is the default data provider for MetaTrader 4, and it provides access to a wide range of financial instruments from various exchanges, including FOREX.
- FXCM: This is another popular data provider for MetaTrader 4, and it offers access to a variety of currency pairs and other financial instruments.
- OANDA: This is a well-known provider of forex and CFD trading services, and it offers data feeds for MetaTrader 4 that include real-time and historical data.
- Dukascopy: This is a Swiss-based forex broker that offers data feeds for MetaTrader 4, including tick data and historical data going back several years.

API's

In addition to these providers, there are also APIs available that can provide you with access to exchange data. For example, the OANDA API allows you to retrieve real-time and historical exchange rates for a variety of currency pairs. Other APIs that provide similar functionality include:

- Polygon.io Currencies package API: free version with: 5 calls / minute, 2 years historical data, end of day data
- XE Currency Data API: free trial
- Open Exchange Rates API: free trial

While some of these options offer free trials, every option eventually charges the user, except for MetaQuotes.

MetaQuotes is the default data provider for MetaTrader 4, and the platform comes pre-installed with it. However, while the MT4 platform itself is free to download and use, accessing real-time market data and other advanced features may require a subscription or a fee. The real time market data is something that can be included in a providers feed. While having an external source is nice, **we do not really need any actual live data.**

Import historical data in local database

There are several free data providers and APIs that offer historical data for a wide range of financial instruments, including FOREX. Here are some options that provide access to extensive historical data:

- Dukascopy: Offers free access to tick data going back several years for a limited selection of currency pairs. The data can be downloaded in a variety of formats, including CSV and JSON.
- TrueFX: Offers free access to historical tick-by-tick data going back several years for a selection of currency pairs. The data can be downloaded in CSV format.
- HistData: Offers free access to historical data for a variety of financial instruments, including FOREX, going back several years. The data can be downloaded in CSV format.
- Forexite: Offers free access to historical data for a variety of currency pairs going back several years. The data can be downloaded in CSV format.
- Alpha Vantage: Offers free access to historical data for a variety of financial instruments, including FOREX, going back up to 20 years. The data can be accessed through an API and downloaded in various formats, including CSV and JSON.

It's important to note that free data providers may have limitations on the amount of data available or the frequency of updates, and the quality of the data may vary.

The services above allow us to download the historical data (mostly) as csv files. According to documentation of MT4, you can import historical data in CSV format using the platform's History Center feature. Here are the steps to import historical data into MT4:

1. Download the historical data in CSV format from your chosen data provider or source.
2. Open the MT4 platform and go to the "Tools" menu and select "History Center".
3. In the History Center window, select the financial instrument (e.g. currency pair) for which you want to import historical data.
4. Click on the "Import" button at the bottom of the window.
5. In the Import window, select the CSV file you downloaded and click "Open".
6. In the Import settings window, select the appropriate settings for the import, including the date format, time zone, and price format.
7. Click "OK" to start the import process.
8. Once the import is complete, you should see the imported historical data in the History Center window.

It's important to note that the imported historical data will only be available for use in MT4 once it has been added to the platform's database, which may take several minutes or longer depending on the size of the data set. Additionally, not all CSV data formats may be compatible with MT4, so it's important to check the data format requirements for the platform before importing the data.

The functionality of importing data should also be possible using the command line.

Using imported data

Once the data is imported, you can view it in various ways within MT4, such as using the charting tools or running backtests on your trading strategies. The data will also be used by any Expert Advisors or other custom indicators that you may have installed.

It's worth noting that MetaTrader 4 has limitations on the amount of historical data that can be stored in its database. By default, the platform only stores a limited amount of historical data, and older data may be automatically deleted over time. However, you can adjust the settings to increase the amount of data that is stored, or you can store the data externally and load it into MT4 as needed.

2. Are there any differences in the data provided by different sources?

To answer this question, I used **literature study** in order to identify all of the ways exchange data can differ from each other. Below are the main ones important for the project:

Different sets of currencies

Not all providers provide all currency sets. This would mean that with certain data providers, not all currency pairs could be tested on.

Internet connectivity and Request / response via HTTP or Web sockets

With some data sources, the user may only download the data manually. Other have the option to request it using their services. For the sources that provide this, the way the request / response is given can also differ. Some services offer Req./Responses via REST/HTTP and others use web sockets. This difference won't matter as the application won't run strategy tests on live data.

Data frequency

When a source delivers data, it delivers price changes of currency pairs. Providers can differ in the frequency of these price changes. For the project, it's better to have the highest frequency possible. This prevent the faulty closing of positions hitting their target / stoploss during strategy testing.

Data time period

Forex data always has a timestamp. The way the sources discussed earlier often deliver their data is by packaging all of the data of a certain currency pair, for a certain month. For example, they'll package all of their data of EUR/USD for May of 2022. This means the sources differ from each other because they provide different packages. Some sources only give away the last 24 months, while other may only give random forex currency pair data for certain months between the years 2000 and 2010.

Tick Data Type

Tick data has 2 standards it can follow. The first standard has a bid and ask value. The second standard has a price with a buy/sell Boolean. Both standards give the same amount of information. If the project were to include the creation of an exchange information API, it would be beneficial to pick one of these standards, and implement them into the project of all groups working on the project.

For this project, we want as many currency sets as possible. The data also need to have a frequency (resolution) of milliseconds. The tick data type has two different standards it can follow. Both give the same information but we should choose one in order to make the application compatible with the exchange data source. This standard should also be communicated with the other group working on this project in order to simplify any possible future collaborations.

To easily compare all the data options, I used Multi-criteria decision making and made a table. For all data below, it was assumed the user would use the free version / trial of the data sourcing application.

	Sets of currencies	HTTP / Web / Download	Data frequency	Time period	Tick data type
Dukascopy	Major, minor and exotic pairs	Limited http API, download	Historical: 1m-5m real time: 1s	2003 – now, limited number of months per currency pair	both
TrueFX	Major pairs	download	Tick by tick, not exactly specified	2009 – now, limited number of months	both
HistData	Major pairs	Download via http requests, http API	Historical: 5m, 15, 1h	2001 – now, limited number of months	both
Forexite	Major and minor pairs	Limited download	Daily and hourly	2001 – now, limited number of months	both
Alpha Vantage	Major pairs	Historical and realtime http API,	API: 1m – 1d (higher frequency may need subscription)	1999 – now, real time is current trading day only, most historical data requires subscription	both

3. How do you evaluate the quality of data provided by each source?

This section aims to provide a framework for evaluating the quality of data from different sources, including primary and secondary sources. The section discusses various factors that can affect the quality of data

Data completeness

The first factor to consider is whether the data provided by each source is complete. Incomplete data can lead to biased conclusions, and it may be necessary to supplement it with additional sources or techniques to fill in gaps.

Data relevance

The second factor is whether the data is relevant to the research question or objective. Data that is not relevant may introduce noise or confusion to the analysis and detract from the quality of the results.

Data accuracy

The third factor is the accuracy of the data provided by each source. Accuracy refers to how closely the data matches the real-world phenomenon or event being measured. This can be evaluated by comparing the data to known or validated sources, or through statistical methods.

Data consistency

The fourth factor is whether the data is consistent over time or across different sources. Inconsistencies can be a sign of errors or inconsistencies in the data collection process.

The factors above can be used in the future to compare and evaluate the quality of different data providers, but for the time being, according to the stakeholder, the quality itself isn't important yet. The focus now is to find a source with sufficient quantity.

4. How accurate and reliable is the data provided by each source?

This section aims to provide guidance on how to assess the accuracy and reliability of data provided by each source. The section discusses various methods that researchers can use to verify the accuracy and reliability of data

Data source credibility

The first factor to consider is the credibility of the data source. This can be evaluated by looking at the reputation of the source, the qualifications of the people collecting the data, and any biases that may be present

Data collection methodology:

The second factor is the methodology used to collect the data. Data that is collected using standardized methods or procedures is generally more reliable and accurate than data that is collected using ad hoc or informal methods.

Data sampling

The third factor is the sampling method used to collect the data. A sample that is representative of the population being studied is more likely to be accurate and reliable than a sample that is biased or incomplete.

Data validation

The fourth factor is whether the data has been validated or verified through independent sources or methods. Validation helps to ensure that the data is accurate and reliable and can improve the quality of the analysis and conclusions.

The factors above can be used in the future to compare and evaluate the accuracy and reliability of different data providers, but for the time being, according to the stakeholder, these factors aren't important yet. The focus now is to find a source with sufficient quantity.

Conclusion

When taking into consideration that quality and reliability aren't high priority for the time being, it's up to the amount of available data. When comparing all the sources in sub-question 2, the best option seems to be TrueFX. This is because It's the only source with a tick frequency that's high enough to meet our stakeholder's wishes. It also provides all of the major currency pairs.

Sources:

Dukascopy:

https://www.dukascopy.com/land/trading/swfx/eu/forex/?lang=en&gad=1&gclid=CjwKCAjw0ZiiBhBKEiwA4PT9z230g2Am8SDBOl3rpgGJqyybSJzJ2i1A7774Gu65NGP43EZQKARzUBoCl9QQAvD_BwE

Forexite:

https://www.forexite.com/traderoom/

Alpha Vantage:

https://www.alphavantage.co/

Polygon.io:

https://polygon.io/dashboard/currencies

TrueFX:

https://www.truefx.com

HistData:

https://www.histdata.com/download-free-forex-data

Data quality info:

https://www.cloudresearch.com/resources/guides/ultimate-guide-to-survey-data-quality/guide-data-quality-what-is-data-quality-why-important/