Instituto BME

A logo with blue and grey letters

Description automatically generated

Master of Science Thesis

Optimizing Global Portfolio Weights using Artificial Intelligence and Macroeconomic Analysis

Master in Artificial Intelligence and Quantum Computing Applied to Financial Markets, 11th edition (mIA-X)

Lukasz Tadeusz Harazin

July 2024

Table of Contents

[1. Introduction 3](#_Toc165808807)

[2. Background 4](#_Toc165808808)

[3. Methodology 5](#_Toc165808809)

[1.1 Data sources 5](#_Toc165808810)

[1.1.1 Organisation for Economic Co-operation and Development (OECD) 5](#_Toc165808811)

[1.1.2 World Bank (WB) 5](#_Toc165808812)

[1.1.3 Bank of International Settlements (BIS) 6](#_Toc165808813)

[1.1.4 Yahoo Finance 6](#_Toc165808814)

[1.1.5 Investing.com 6](#_Toc165808815)

[1.1.6 Morgan Stanley Capital International (MSCI) 7](#_Toc165808816)

[1.1.7 Other sources 7](#_Toc165808817)

[1.2 Macroeconomic Data 8](#_Toc165808818)

[1.3 ETF Data 8](#_Toc165808819)

[4. Presentation of work 9](#_Toc165808820)

[5. Conclusions 10](#_Toc165808821)

[6. References 11](#_Toc165808822)

# Introduction

TODO:

* Problem definition
  + describe what you trying to do
  + clearly state the question being addressed
  + when appropriate formulate a testable hypothesis
* Describe the motivation; who is interested in the solution.
* Summarize the results and their significance.

# Background

TODO:

* Describe current understanding of the problem, existing solutions, and the barriers to these solutions.
* Review of the pertinent literature.

# Methodology

## Data sources

In my master thesis I have relied exclusively on open databases available on the internet. These sources might not be as good as paid data providers like Bloomberg terminal, being the gold standard, but still provide sufficient data coverage to perform detailed macroeconomic analysis.

There are various international organizations collecting broad range of macroeconomic statistics which are publicly available. Main organizations with open datasets are World Bank (WB), Organisation for Economic Co-operation and Development (OECD), International Monetary Fund (IMF), United Nations (UN) and Bank for International Settlements (BIS).

Additionally, there are multiple financial websites that offer a variety of resources for investors and traders. The most popular portals providing market data and economic indicators are Investing.com, Yahoo Finance, Trading Economics and FX Empire.

After extensive analysis of all available data sources, I have selected the best sources with the highest data coverage for all countries selected for global portfolio. Below, I will describe all data sources used in my algorithm.

### Organisation for Economic Co-operation and Development (OECD)

The Organisation for Economic Co-operation and Development (OECD) compiles a comprehensive collection of economic, social, and environmental datasets encompassing member countries and selected non-member economies. These datasets include:

* National accounts, trade, labour, and productivity statistics
* Science and technology innovation indicators
* Entrepreneurship and formation data
* Environmental-economic accounting information
* Development resource flows and official development assistance (ODA) data

In my macroeconomic analysis, I have used a wide range of monthly and quarterly economic indicators. It has the best data coverage for most countries included in MSCI ACWI index. Data can be downloaded using csv exports or developer’s API. In my algorithm, I have managed to use multiple indicators from OECD.Stat website reaching back to 1999.

### World Bank (WB)

World Bank Data is a comprehensive resource for global development data. It offers data by country or indicator, databases, pre-formatted tables, reports, and other resources. It includes tools for data analysis and visualization, access to microdata, international debt statistics, and development indicators. It also provides information on World Bank’s finances and lending projects.

For my analysis, it includes even more countries than OECD. Unfortunately, it provides mainly yearly data with significant delay. That limits its use in terms of monthly portfolio rebalancing. Data can be exported using csv files or third-party python libraries providing an easy access to all indicators formatted automatically to pandas’ data frames.

### Bank of International Settlements (BIS)

The BIS Data Portal provides global financial statistics. It includes data on international banking activity, debt securities, credit to non-financial sectors, credit-to-GDP gaps, debt service ratios, global liquidity indicators, exchange-traded derivatives, OTC derivatives, triennial survey data, and residential property prices. All data can be downloaded in CSV or SDMX format. Moreover, databases can be accessed via REST API providing data in XML format.

In my work I have used its data for central bank rates. I haven’t found any other place with such a complete daily data for interest rates for almost all central banks in both developed and emerging economies. Data is updated weekly what provided very precise and up to date information.

### Yahoo Finance

Yahoo Finance is a financial website that provides investors with:

* Market Data: Real-time stock quotes, news, and analysis.
* Investment Tools: Portfolio management tools and research resources.
* Financial News: Up-to-date headlines and insights on global markets.

Data can be easily downloaded using open-source Python library which reads real time market data from publicly available APIs. I have used this library to read historical prices for ETFs, stock indices and currency rates.

### Investing.com

Investing.com is a leading financial website and platform that offers a variety of tools and resources for investors and traders. It offers:

* Real-time market data: Stocks, bonds, commodities, currencies, futures, options across 70+ exchanges.
* Analysis & news: Articles on market trends, company performance, and economic events.
* Investment tools: Economic calendar, earnings calendar, technical analysis tools, portfolio tracker.

In my work, its economic calendar has provided a great value. While it has slightly worse data coverage than OECD dataset, it does provide report date time for all indicators. That is immensely important for proper back testing because we know exactly what data was available at a certain point in time. While other databases only hold final revision for given indicators like GDP values, with exact investing calendar we can read all values for subsequent revisions and know on which days they were released. Economic calendar cannot be easily downloaded, so I have used techniques of web scrapping using Selenium library. I have managed to download over 320 thousand data points for 50 countries in MSCI ACWI index dating back to 1999.

Apart from economic calendar, I have also downloaded csv data for missing currency rates and selected stock indices that aren’t available on Yahoo Finance. Moreover, it has complete historical data for Manufacturing PMI indicators, hard to find anywhere else, which are very useful to calculate economic cycles.

### Morgan Stanley Capital International (MSCI)

Morgan Stanley Capital International maintains a family of stock market indexes which are widely followed by investors around the world. MSCI indexes are widely used benchmarks for global stock markets. They track different segments (like developed, emerging or country-specific) and by market cap (where bigger companies influence more). This allows investors to see how their portfolios perform compared to a specific market segment.

In my algorithm, I have used MSCI indexes to both obtain MSCI ACWI benchmark and also fill missing returns on selected exchange traded funds where Yahoo Finance doesn’t have complete data.

### Other sources

Apart from data sources listed above, I have explored other datasets and web portals:

* **The International Monetary Fund (IMF)** Data provides comprehensive economic, financial, and socio-demographic statistics. It includes data on direct investment, climate transition, greenhouse gas emissions, world economic outlook, international finance, global financial stability, fiscal monitor, and exchange rates. It covers many similar indicators to OECD, however it has worse historical data coverage and more missing values for old data. In the end I opted to use OECD datasets instead.
* **Trading Economics** is a platform that provides similar features to Investing.com, but is better protected against automatic web scrapping and offers paid subscriptions to access its data.
* **FX Empire** is another data provider resembling Investing.com and Trading Economics. Unfortunately, its interactive website makes it very difficult to download data using web scrapping.
* **FRED** (Federal Reserve Economic Data) is a trusted source for economic data since 19911. It provides access to over 824,000 US and international time series. The data covers various economic indicators such as CPI, GDP, inflation, unemployment, and exchange rates. It does provide an excellent coverage for US economy, but lacks international indicators, referencing data from OECD and other public database.
* **EBS Statistics** provide comprehensive data that supports all aspects of the ECB’s work, including monetary policy, financial stability, and banking supervision. It covers European economies in great details but covers few international indicators required to optimise global portfolio.

## Macroeconomic Data

## ETF Data

# Presentation of work

TODO:

* Describe models and results

# Conclusions

TODO:

* Summary of results
* Recommendations: generalize conclusions to appropriate design decisions, practices and/or procedures
* Implications for further study
* Future Work

# References

TODO:

* List of references