第一页为封面页

参赛队员姓名：	闫翔宇
中学：	西安高新第一中学
省份：	陕西省
国家/地区：	中国
指导教师姓名：	张恩典
指导教师单位：	西安高新第一中学
论文题目：	Portfolio research based on Climate Index and ESG rating

第二页此处开始为论文的主体部分……

论文题目：Portfolio research based on Climate Index and ESG rating

作者：闫翔宇

摘 要： In recent years, the global climate has changed dramatically, and various extreme weather events have occurred frequently. While it has an impact on activities such as daily life, production and operation, it also has a wide-ranging impact on the development of economy and finance. ESG is a comprehensive consideration of environmental, social and corporate governance factors, with more emphasis on the sustainable development of enterprises. And, when it comes to combating systemic risks, such as Covid-19 and climate change, companies with well-established ESG management systems have clear advantages.

Based on the climate index and ESG rating, this paper uses the optimized model to construct the investment portfolio, and uses the climate index tracking strategy and the interval channel strategy for backtesting. First, according to the climate change white paper, an authoritative climate-related text corpus was constructed, including key words such as temperature, sea level, and isoprene glycol. Using the tf-idf algorithm, according to the frequency and similarity of keywords, a continuous climate risk index is synthesized, and the final climate index is obtained after smoothing. Second, set a suitable threshold for ESG, only stocks above the ESG threshold can be included in the portfolio. The portfolio is then constructed by optimizing the model and solved in Python. Finally, backtesting analysis with different strategies shows that the portfolio constructed in this paper is effective and performs well in terms of risk control and returns.

关键词：Climate Index; ESG; Portfolio; Nonlinear Programming

Table of Contents

1. Introduction

1.1 Background

There has been significant climate changes in our country in recent years, with extreme weather conditions occurring frequently. In the winter of 2021, there has been large scale extreme cold weather in the northern provinces, and in the summer of 2022, Chongqing has seen temperature as high as 45 degrees Celsius, all of which have had serious impact on people’s lives. In June 2022, 17 government departments, including the Ministry of Ecology and Environment, the National Development and Reform Commission, and the Ministry of Finance, jointly issued the “National Climate Change Adaptation Strategy 2035”. This made overarching plans and arrangements for the work required for our country between now and 2035 to adapt to climate changes, and pointed out that such work should be integrated into the overall economic and social development. In addition, in October 2020, five departments, including the People's Bank of China and the China Banking and Insurance Regulatory Commission also issued the "Guiding Opinions on Promoting Investment and Financing to Address Climate Change" to guide and leverage more social funds into the field of climate change. It can be seen that the drastic changes in climate not only affect daily life, but also affect national policy formulation and economic and financial development.

ESG ratings are produced by professional rating agencies, evaluating a company from the perspectives of environment, social responsibility and corporate governance. In recent years, both domestic and international community have paid growing attention to ESG ratings. Buying stocks of companies with high ESG ratings is conducive to selecting truly high-quality sustainable companies for investment, and is conducive to protecting the environment and improving climate conditions.

Climate change affect the operations of agriculture, industry, finance, and technology companies, and as a result the stock market. However, whether climate factors can be used as a basis for constructing investment portfolios, and the effectiveness of such portfolios, are both subjects of continued research. In this paper I firstly quantify the climate changes by constructing a climate index. Secondly, I construct a portfolio, which is based on the CSI 300 constituent stock universe. Such portfolio is constructed with the joint objectives of maximizing returns and minimizing tracking errors with respect to the climate index, with the ESG scores as constraints, filtering stocks based on their ESG scores. I have then set a threshold value, buying the portfolio when the index is above the threshold, and selling the portfolio vice versa. Finally, I have ran various backtests to analyze the performance of my trading strategy.

1.2 Recent research

In terms of climate index construction, existing research can be divided into two categories according to the type of data studied. One is to directly analyze meteorological observation data. Sea level pressure and other meteorological factors are used to construct a static and stable climate index by dividing the threshold range by historical sample statistics and then summing the weights. The second is to analyze text data, and indirectly obtain weather-related data through weather reports such as newspapers and periodicals. Engle R F (2019) used the tf-idf statistical method to construct a climate index that can provide a reference for investment decisions by counting the frequency of weather-related keywords in financial newspapers. TF-IDF is a statistical method to assess the importance of a word to a document set or a document in a corpus. The importance of a word increases proportionally to the number of times it appears in the document, but decreases inversely to the frequency it appears in the corpus. Since the direct meteorological observation data cannot be fully reflected in the stock market, and the content of financial newspapers and periodicals is much more closely referenced by stock market participants, I adopt the second approach to construct the climate index by utilizing the weather information in financial newspapers and periodicals.

On the subject of index outperformance, a common method is to use an optimization model to construct the investment portfolio with minimum tracking error and maximum expected return, subject to various constraints, so as to obtain excess returns higher than the average return of the index while tracking the trend of the index. Huang Jinbo (2019) used partial moment as a new constraint to construct an index enhanced model, which can control tail risk and obtain excess returns. Zhao Zhihua (2016) introduced a perturbed set of returns, established a sparse robust optimization model, and confirmed that the model significantly reduced the tracking error while ensuring the out-of-sample excess returns. In this paper, I employ a nonlinear optimization model to construct portfolios that seek to outperform the climate index.

1.3 Research design

My research steps are shown in Fig1-1.

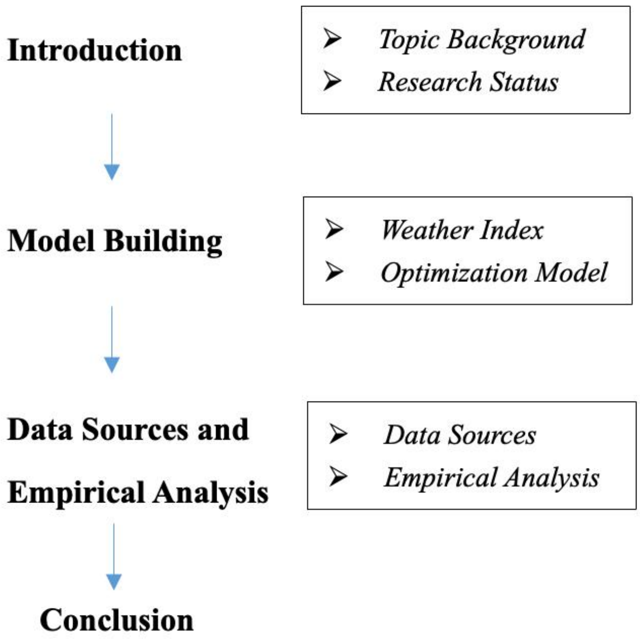

Fig1-1 Research steps

As shown in Figure 1-1, the first part of this paper is the introduction, including the background and recent studies. Based on the research background and recent studies, I have determined the research direction and method. That is, firstly construct the climate index by analyzing text data, and then construct the investment portfolio via nonlinear optimization method. The second part of this paper describes the specific construction method of the climate index and the optimization model. The third part of this paper is an empirical analysis. Based on historical data, I construct the climate index, solve an optimization problem to obtain an investment portfolio, and finally analyze the return characteristics of the portfolio. The fourth part of this paper is the conclusion, which summarizes the research results.

2. Model construction

2.1 The climate index

Overview: I count the frequency of climate keywords in daily news, and compare the frequency of climate keywords in standard white papers on climate issues, to arrive at a measure of daily climate riskiness. Each day's climate risk measure are then combined into a continuous climate index that reflects the overall changes in climate risk.

The rationale of this approach is that, when there is a climate risk issue, relevant news reports will increase accordingly, and the increased number of relevant reports indicates higher climate risk. Firstly, I have identified 128 "climate keywords" through various related text data, such keywords include weather, earth surface, and so on. For the climate news, I have selected Securities Times Electronic News, Shanghai Securities Electronic News and Securities Daily Electronic News. These three news sources are selected because they are all subsidiaries of the People's Daily, selected by the China Securities Regulatory Commission, the China Insurance Regulatory Commission and regulatory bodies as the designated disclosure platforms, with extensive information coverage and high authority, and the news texts are easy to obtain. Another reason is that the main research direction of this paper is the relationship between climate risk and the market, and the news sources that individuals and institutional investors pay more attention to in the Chinese market have a greater impact on the market. After identifying these news sources, I have downloaded all the raw news text data for a total of three years from 2019 to date. Then for every day I calculated the total size of the texts as well as the frequencies of each climate keywords in those texts.

In addition, in order to obtain standard vocabulary on climate issues, a total of 8 white papers on China's policies and actions on climate change from 2013 to 2021 were selected as a comparison. The eight articles include: "White Paper on Forestry's Policy and Action on Climate Change in 2013", "White Paper on Forestry's Policy and Action on Climate Change in 2014", "White Paper on Forestry's Policy and Action on Climate Change in 2015", "2016 White Paper on Forestry's Policy and Action on Climate Change" White Paper on Climate Change Policies and Actions, 2017 White Paper on Forestry and Grassland Policies and Actions on Climate Change, 2018 Forestry and Grassland Policies and Actions on Climate Change, and 2019 Forestry and Grassland Policies and Actions on Climate Change , "2021 White Paper on China's Policies and Actions on Climate Change". The eight articles were put together as a standard climate literature, and the frequency of each climate keyword in this article was counted.

The index construction was mainly based on the tf-idf algorithm. This is a simple but efficient algorithm to determine the relevance of each keyword in a corpus. Tf stands for term frequency, which is the frequency of each keyword. The higher the frequency, the more important the word is in this article. Idf stands for inverse document frequency, the idea is that if the word appears in every article, it means that the word is less important. For example, if the word water is mentioned in every article, it means that the appearance of the word water itself cannot bring additional information. The main purpose of this step is to reduce the excessive weight given to common words.

In the calculation step, the tf-idf value corresponding to each climate keyword every day is calculated first using the news data of the electronic newspaper. This represents the importance of the climate keyword on this day, which is denoted as X_{it}, where t is date, and i is the climate keyword. Then, the tf-idf value corresponding to each climate keyword in the white paper is calculated, which represents the importance of the climate keyword in the white paper, denoted as Y_i, where i is again the climate keyword. Then, the cosine similarity between X_{it} and Y_i is calculated, and this value is used as the value of the climate index for that day. Higher similarity value indicates higher news focus on climate related issues, and hence higher climate index values, and vice versa. Finally, the daily climate scores are concatenated together to form a continuous climate index, and then smoothed to become the final climate index, as shown in the following figure.

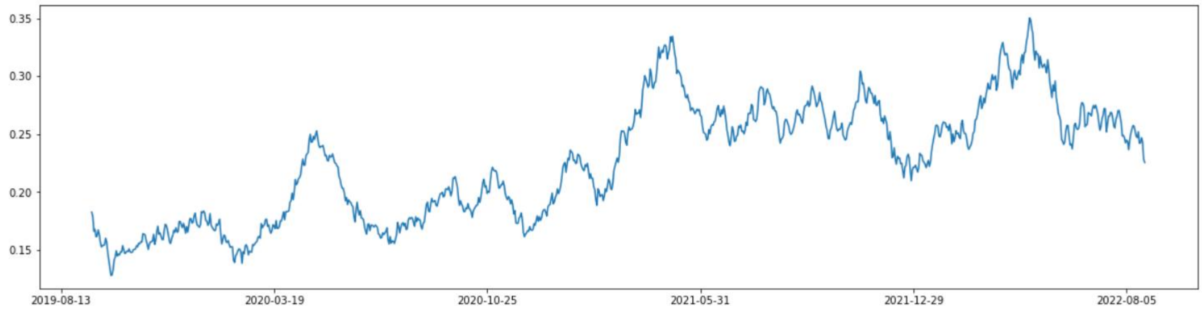

Fig 2-1 Climate index

As can be seen from Figure 2-1, the climate index has generally shown an upward trend in the last three years, which is in line with the prior expectation. In addition, the climate index shows a strong seasonal pattern, with a small peak around March and April every year, which may be related to the news about smog every spring.

One potential problem with this approach is that it doesn't distinguish between positive and negative news. When the news about climate increases, it may indicate that there has been a recent deterioration in climate conditions, but it may also be because of government policies to strengthen governance, or news reports about the environment getting better. These different types of news can have significantly different impacts on the market. Therefore, a future research direction of this paper is to subdivide news sentiment to optimize the calculation of climate risk.

2.2 Model optimization

I construct a portfolio with the help of an optimization model, assuming that there are N stocks in the portfolio and a total of M periods / observations. On the one hand, smaller the M-period average tracking error between the portfolio and the climate index is preferred, so the objective function is as shown in Equation 2.1:

$$\min_{\omega} TE = \frac{1}{M}\left\| R\omega - I_{RETURN} \right\|_2^2 \qquad (2.1)$$

Here, R is the returns matrix, $R \in \Re^{M \times N}$; I_{return} is the climate return index, $I_{return} \in \Re^{M}$; ω is the stock weights vector, $\omega \in \Re^{N}$; TE is the tracking error between the portfolio and the climate index。

On the other hand, a suitable threshold needs to be chosen, so that when the climate index is lower than the threshold, it means that the climate risk is low. At this time, we should long the investment portfolio, and the M-period average excess return of the investment portfolio should be as large as possible. The corresponding objective function is as shown in formula 2.2

$$\max_{\omega} ER = \frac{1}{M} e^T (R\omega - I_{RETURN}) \tag{2.2}$$

Here，e is the unit vector，$e \in \Re^M$，ER is the mean excess return of the portfolio over the climate index.

Merging eq 2.1 and 2.2 to a single objective function, as seen in eq 2.3:

$$\min_{\omega} ER = \tau TE - (1-\tau) ER \tag{2.3}$$

Here τ is the relative weights of the two objectives，which can be configured to place different weights on the objectives. $\tau \in (0,1)$。

The constraints of the model can be found in eq 2.4:

$$\begin{aligned} &\mu^T \omega \geq \varepsilon_{ESG} \\ &e^T \omega = 1 \\ &\|\omega\|_0 \leq K \end{aligned} \tag{2.4}$$

Here，μ are the ESG scores，$\mu \in \Re^N$；ε_{ESG} is the minimum ESG required；$\mu^T \omega \geq \varepsilon_{ESG}$ indicates the minimum requirement for ESG scores. $e^T \omega = 1$ indicates the stock weights should sum to 1. $\|\omega\|_0 \leq K$ means there are K stocks at most.

Now I need to solve the model using scientific library scipy in Python, passing in four sets of parameters: initial values, boundaries, constraints, and objective function. I use the built-in functions to solve the restricted optimization problem with constraints. After 50,000 iterations, the optimal result is found and the results are exported into a table. Each row of the table is a date, each column is a stock code, and the value of the corresponding cell is a stock weight, that is, the portfolio under this set of weights can satisfy the constraints and minimize the value of the objective function.

When the climate index is above a predetermined threshold, I should in theory build another portfolio that underperforms the index as the climate risk is perceived to be too high. But due to the constraints of shorting individual stocks, I short portfolio instead.

3. Empirical analysis

3.1 Data source

The data used in this paper come from climate journals, securities newspapers, and the Wind database. Among them, climate journals include the "Climate White Papers" in the database of CNKI (https://www.cnki.net/): "China's Policies and Actions to Address Climate Change", "Future Weather and Climate Prediction", "Enhanced Climate Change" Action on Climate Change - China's Nationally Determined Contribution, etc. Securities newspapers include: (1) Securities Daily Electronic News: http://epaper.zqrb.cn/; (2) Securities Times Electronic News: http://epaper.stcn.com/; (3) Shanghai Securities Electronic Newspaper: https://paper.cnstock.com/.

3.2 Data retrieval

The above raw data are all on the web page in HTML format. I have obtained these data using web crawler technology in Python. Among them, the core technologies used include:

- The development language is in Python.
- The requests tool is used to send http requests to download the HTML source code from the website.
- The lxml tool is used to parse the HTML source code into DOM (Document Object Model) tree.
- The XPath technology is used to obtain the required data in the DOM tree, such as article title, body URL, body content.

Specific steps are as follows:

- Find the website containing the data needed from the search engine.
- Find the daily e-newsletter directory URL from the web page, and analyze the URL rules.
- Recursively obtain the daily raw data, including title and text URL separately.
- Visit the text URL to obtain the text of the webpage (remove non-text content such as pictures and videos).
- Save the acquired data to a txt format file on the local hard drive.

3.3 Data processing

First, in order to quantify the intensity of climate news reports in securities newspapers, an authoritative text corpus of climate topics was constructed from more than 10 climate change white papers, and natural language processing technology was used to aggregate them into 114 climate change keywords in Chinese and English, for example: temperature, sea level, isoprene glycol, ozone, basal radiation, methane, emissions, etc.

Next, I use regular expressions to remove spaces and punctuation marks in the body of the newspaper, leaving only Chinese, English, and numbers. The total number of words in the text is counted, and the number of times each keyword appears in the text is searched in turn. If no keyword appears, the number of times is counted as 0. I then write the statistical results such as date, title, source website, total word count, and occurrences of each keyword into a csv file.

Finally, I downloaded all the stock returns of the CSI 300 and their FTSE Russell ESG scores in the corresponding time period from the Wind database. During certain time periods, the ESG scores of certain companies will be missing, and the scores will be filled in with the ESG ratings of SynTao Green Finance. For companies with neither ESG Scores nor SynTao Green Finance Ratings, we removed them from the stock universe. After processing, the final stock universe contains 293 stocks of CSI 300.

3.4 Backtest settings

I have used the JointQuant platform for backtesting. The backtesting interval is from September 27, 2019 to August 19, 2022. The starting capital is set to 5 million yuan, with stock prices dynamically adjusted for corporate actions. The transaction fees are set at 3bps commission for both buying and selling, and an additional 10bps of stamp duty when selling, with a minimum commission for each transaction is 5 yuan. The specific backtesting ideas are as follows:

- Calculate individual stock weights based on the model, rebalancing on daily basis.
- Use 3 configurable parameters including the trade-off parameter $\tau \in (0,1)$ minimum ESG score ε_{ESG}, and the upper limit of stock holdings K。

3.5 Backtesting

3.5.1 Climate index tracking strategy

First, I only consider a simple index tracking strategy. I construct long only investment portfolios with similar performance as the climate index, and ignore short selling. Under the climate index tracking strategy, the choice of trading frequency is an important issue. It is more common to set the transaction frequency to 1 day, 3 days, 5 days, 21 days, and 30 days. Therefore, the excess returns enhanced by the aforementioned 5 trading frequencies are used to backtest the index, and the specific returns are shown in the figure:

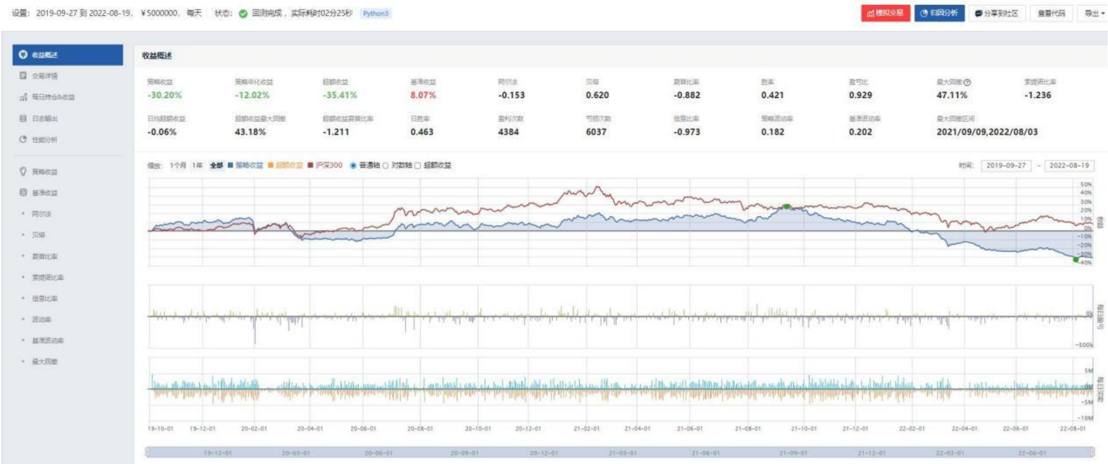

Fig 3-1 Excess performance with 1 day rebalancing

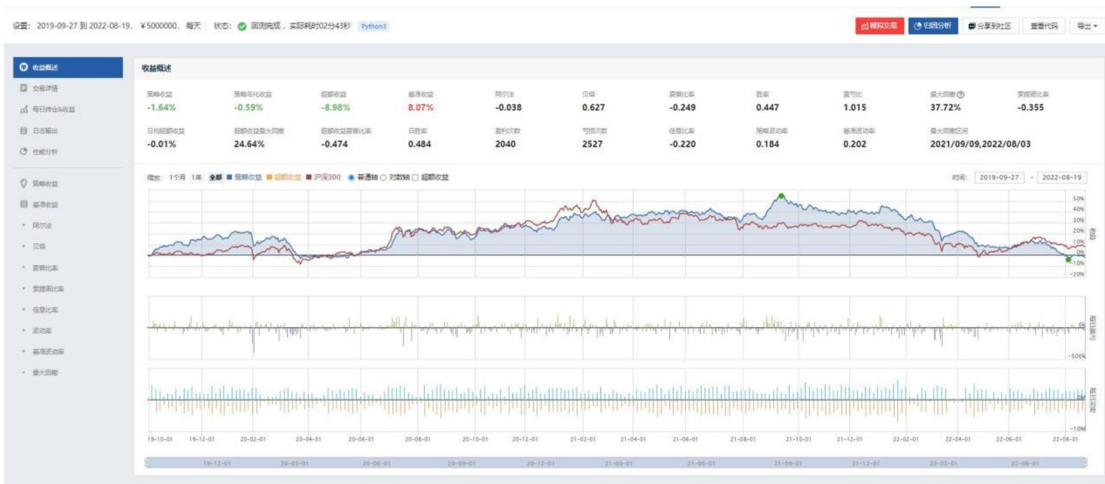

Fig 3-2 Excess performance with 3 day rebalancing

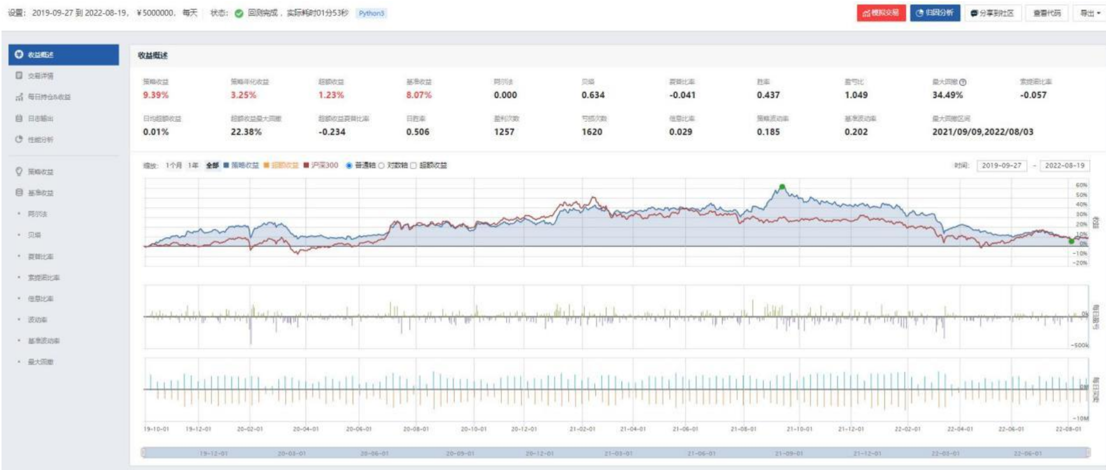

Fig 3-3 Excess performance with 5 day rebalancing

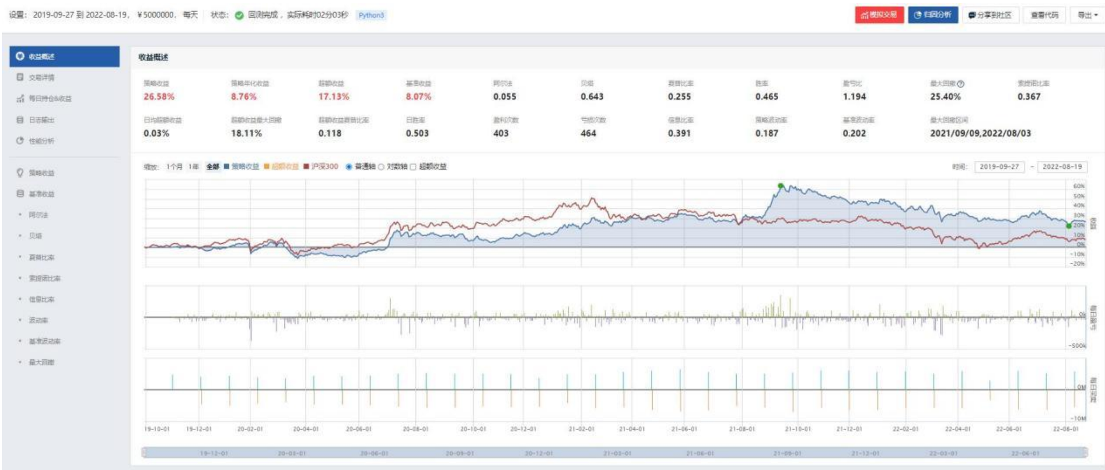

Fig 3-4 Excess performance with 21 day rebalancing

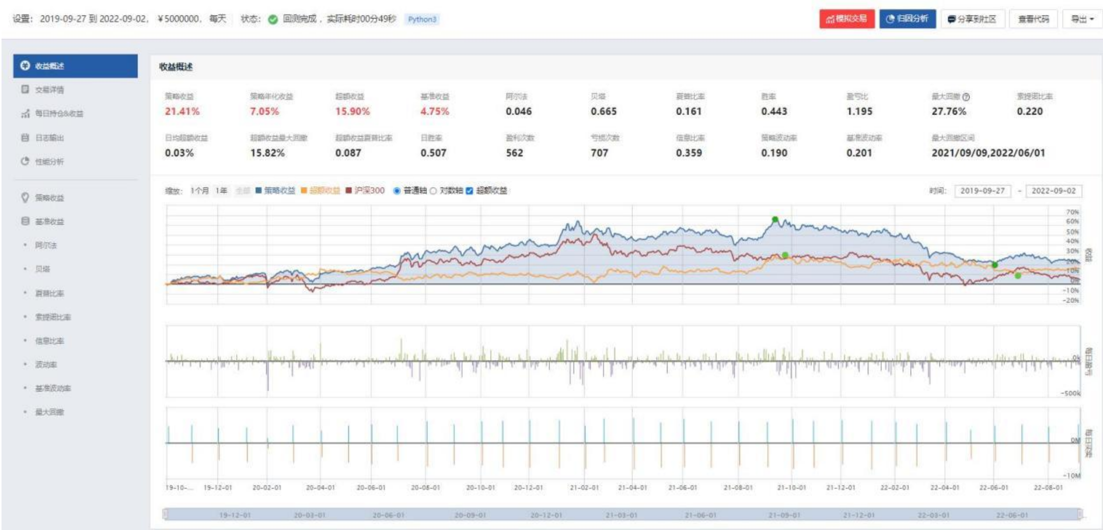

Fig 3-5 Excess performance with 30 day rebalancing

As can be seen above, the excess return of the strategy varies with the trading frequency. When the trading frequency is too high, the noise is high and the amount of information is low, which makes it impossible to effectively capture the information contained in the index, resulting in a low excess return, and even in some cases, unable to outperform the return of the CSI 300. However, when the trading frequency is too low, the trend is more obvious, which will also affect the overall excess return of the strategy. Therefore, a climate index tracking strategy is constructed using 21 days as the trading frequency of the strategy.

From the backtest results, it can be seen that the overall return of the strategy is relatively good. During the three-year backtest, the total return is 21.41%, or 7.05% p.a. and the strategy's return during the backtest all exceeded the market, with an excess return of 15.9%. The volatility of strategy returns is similar to that of the broader market, but when the market returns increase, the strategy returns more, and when the market returns decreases, the strategy returns to a smaller extent, which shows that this model performs very well in terms of return and risk control, achieving the strategy goals.

3.5.2 Interval band strategy

3.5.2.1 Basic interval band strategy

The aforementioned climate index tracking strategy has achieved considerable excess returns, but it was a cross sectional strategy with no timing signals. On the basis of the climate index tracking strategy, I construct a basic interval band strategy backtest the result. After the index tracking portfolio is constructed, the returns are further smoothed. I then design a trading strategy based on the smoothed index price in the following way:

- Calculate the upper pressure line and lower support line for the smoothed index.
- Buy when the price crosses the lower support line.
- Due to short selling constraints on A shares, when the price crosses the pressure line, close out the stock positions, and wait until the price crosses the support line again to buy the position again.

The backtest results of the basic interval channel strategy are shown in the figure:

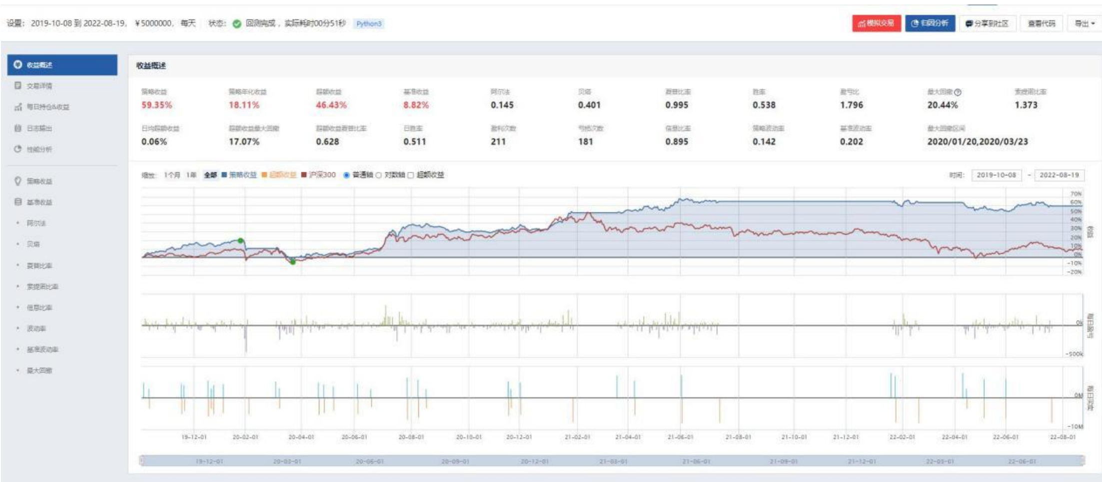

Fig3-6 Basic interval band strategy

Based on the results, the basic interval channel strategy outperformed both the market and the basic index tracking strategy. The total return of the basic interval band strategy has reached 59.35%, and the annualized return is 18.11%. When the market falls, the price goes below the pressure line and the positions are closed. Therefore, after June 1, 2021, the strategy returns are relatively stable; as the market grows, the price break through the lower support line, and the positions are restored, outperforming the market. The Sharpe ratio is 0.995, the information ratio is 0.895, and the winning ratio is 0.538, indicating a good performance in terms of returns vs risk.

3.5.2.2 Improved interval band strategy

Under the basic interval band strategy, due to short constraints, I had to close out all positions as the price falls through the pressure line. However, the futures market has a short-selling mechanism that enables short-selling of CSI 300 stock index futures, that is the IF futures contract. Therefore, I am able to utilize this to improve the basic interval band strategy. When the price crosses the pressure line, I short IF futures, and when the price crosses the support line, the positions are stored as usual. The backtest results of the interval band strategy are shown in the figure:

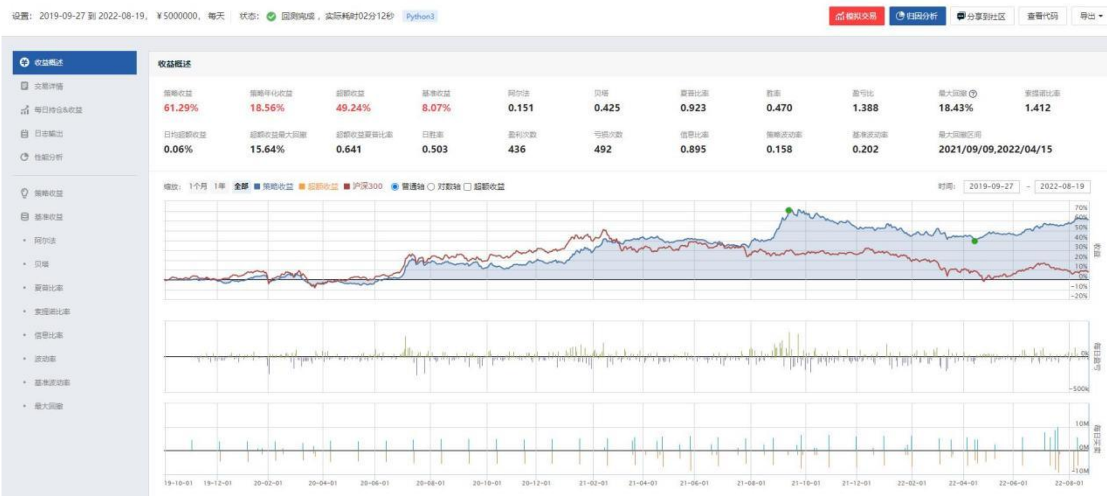

Fig3-7 improved interval band strategy

It can be seen that the performance of the improved interval band strategy is better than the basic interval band strategy, although not significant. The improved interval band strategy overall returned 61.29%, with annualized return of 18.56%, and with an excess return of 49.24%. When the market falls, the price crosses the pressure line and IF futures shorted, which contributed to the overall performance. After June 2021, the strategy return was higher but was also more volatile.

4. Conclusions

（1）In recent years, extreme weathers conditions have occurred frequently, and the global climate problem has become increasingly serious. Based on the climate change white paper, I construct an authoritative climate-related text corpus, including key words such as temperature, sea level, and isoprene glycol, which have received wide attention in recent years. Moreover, the daily climate risk measure is determined according to the frequency and similarity of keywords, and then the daily climate risk degree is concatenated into a continuous climate risk index, and the corresponding smoothing process is performed to obtain the final climate index.

The climate index rose sharply since 2020, and although it has fallen back, the subsequent level is still higher than in 2019. And, starting in 2021, the climate level has once again risen sharply and remains above the previous level. The climate index effectively reflects the increasingly serious climate problems, which serves as a reference for preventing extreme climate.

（2）I constructed a portfolio, taking into account of the ESG scores of individual stocks and demonstrated the effectiveness of the ESG scores. ESG is a comprehensive consideration of environmental, social and corporate governance factors, with emphasis on the sustainable development of enterprises. Therefore, ESG has a long-term impact on the company's goal of maximizing value, and companies with better ESG management systems will have advantages, such as combating systemic risks (Covid-19, climate change, etc.). In addition, ESG rating agencies have comprehensively optimized the rating system and indicator setting, and objectively made ESG ratings and industry benchmarks for companies, so that investors can fully understand the company's ESG governance status, which will promote the value of companies with excellent ESG performance. These factors are linked together to form a positive impact on each other.

Similarly, higher ESG scores will be favored by investors and reduce financing costs, etc., and provide portfolio management with a multi-dimensional perspective on listed companies. Taking ESG as an important perspective, the portfolio constructed in this paper removes stocks below the ESG score threshold, which further illustrates the effectiveness of ESG.

Through empirical analysis, I have demonstrated the effectiveness of portfolio constructed based on climate index and ESG rating. The investment portfolio performs better than the CSI 300 Index, can obtain excess returns far beyond the market, from performance and risk perspective. In addition, comparing the basic interval band strategy and the improved interval band strategy, it can be found that even if the model is improved through the short selling mechanism of the futures market, the backtest performance has not been significantly improved. To a certain extent, it also confirms that China's securities market has a strong bullish sentiment and is not suitable for bearish operations.

I have come to the following recommendations based on the above conclusion:

（1）Investors should be aware of the impact of climate changes on the stock market, and therefore should pay attention to climate related news when making investment decisions.

（2）Sustainable development, environment protection, and preventing extreme climate should be a focus globally.

（3）Investors should also pay attention to the environment and the society, and consider ESG scores as part of the investment decision process.

5. Bibliography

[1]陈懿妮,罗玲,马昊,贾燕,赵军平,徐铖,娄小芬,于燕.杭州市静稳天气指数构建及应用研究[J].环境科学学报,2020,40(12):4461-4469.DOI:10.13671/j.hjkxxb.2020.0232.

[2] Engle R F, Giglio S, Lee H, et al. Hedging Climate Change News[J]. SSRN Electronic Journal, 2019.

[3]黄金波,李仲飞,邹新月.考虑下偏矩约束的增强指数模型[J].管理科学学报,2019,22(12):56-69.

[4]赵志华,徐凤敏,袁晓玲.增强型指数的稀疏鲁棒优化模型及其实证分析[J].数值计算与计算机应用,2016,37(03):199-210.

[5]中国人民银行研究局课题组.气候相关金融风险——基于央行职能的分析.中国人民银行工作论文,2020.

[6]Acemoglu,D.et al.Systemic risk and stability in financial networks, American Economic Review,2015,105(2):564-608.

[7]Acemoglu,D.et al.Transition to clean technology.Journal of Political Economy,2016,124(1):52-104.

[8]Baldauf,M.et al.Does climate change affect real estate prices? Only if you believe in it.Review of Financial Studies2020,33(3):1256-1295.

[9]Barnett,M.et al.Pricing uncertainty induced by climate change.Review of Financial Studies,2020,33(3):1024-1066.

[10]Choi,D.et al.Attention to global warming.Review of Financial Studies,2020,33(3):1112-1145.

[11]Giglio,S.et al.Climate finance.NBER Working Paper,2020b,No.28226.

[12]Pedersen,L.H.et al.Responsible investing: The ESG-efficient frontier.Journal of Financial Economics,2020,forthcoming.

[13]Andersson,M.,P.Bolton,and F. Samama.Hedging climate risk.Financial Analysts Journal,2016,72:13-32.

[14]Balduzzi,P.,and C.Robotti. Mimicking portfolios, economic risk premia, and tests of multi-beta models.Journal of Business and Economic Statistics,2008, 26:354-68.

[15]Hong,H.,F.Li,and J.Xu.Climate risks and market efficiency.Journal of Econometrics,2019,208:265-81

[16]Lamont,O.Economic tracking portfolio.Journal of Econometrics,2001,105: 161-84.

[17]Xuepeng Li,Fengmin Xu,Kui Jing.Robust enhanced indexation with ESG:

An empirical study in the Chinese stock market.Economic Modelling,2021,107 (1):105711.

[18]Zhihua Zhao, Fengmin Xu, Meihua Wang, Chengyi Zhang.A sparse enhanced indexation model with norm and its alternating quadratic penalty method.Journal of the Operational Research Society,2019,70(3):433-445.

Acknowledgements

First of all, I would like to express my high respect and heartfelt thanks to my mentor. From the topic selection of the thesis, the design of the research scheme, the development of the experiment and the processing of the experimental data of this paper, my mentor has given me great help and support. He also spent so much time and energy into guiding my research, and integrated so much wisdom and hard-work for me. His kind and honest conducts, serious and rigorous teaching attitude, and innovatively scientific research thinking have deeply guided and educated me, enabling me to successfully complete this paper.

Secondly, I would like to thank my classmates and my senior alumni for discussing and communicating with me during my study and practice, and for giving me so many valuable opinions and suggestions. In particular, Mr. Tengteng Liu taught me a number of quantitative research methods in the course of my research on stocks. Because of their support and encouragement, I was able to master basic research methods and further set up a lofty academic goal.

Finally, I would like to thank my family. Their selfless dedication is the motivation and confidence for me to successfully complete this paper, and continue to promote my persistent exploration on the academic road.

Appendix – model source code

```
import pandas as pd
import scipy.optimize as sco
import numpy as np

K = 210    # maximum number of stocks with weight 0
CUSTOM_ESG = 1    # parameter ξ esg
    CUSTOM_T  =  0.1    #  parameter τ
INPUT_DIR = r'D:\A_temp_stockW'    # directory for esg.xlsx、index.csv
OUTPUT_FILE = r'D:\A_temp_stockW\stocks_weight_1.csv'    # output directory

N = 0    # For logging, do not change
options = {'maxiter':500000, 'disp':True}    # max number of iterations

def get_weights(returns, index_returns, esg):
    number_of_stocks = len(returns.columns)
    weights0 = [1 / number_of_stocks] * number_of_stocks
    bnds = tuple((0, 1) for _ in range(number_of_stocks))
    #  约束：eq ==0 ineq>=0
    cons = (
                {'type':'eq', 'fun':constraint2},    # Constraint on sum of stock weights
                {'type':'ineq',  'fun':constraint1,  "args":[esg]},     #  Constraint  on  minimum  ESG
score
                {'type':'ineq', 'fun': constraint3 },    # Constraint on number of stocks with 0 weight
    stock_weights    =    sco.minimize(min_fun,    weights0,    args=(returns,    index_returns),
method='SLSQP', bounds=bnds, constraints=cons, callback=callback, options=options)
    data = {}
    for market, weight in zip(returns.columns, stock_weights.x):
        data[market] = round(weight, 6)
        print(market, ':', round(weight, 6))
    return data

def callback(*args):
    global N
    N = N + 1
    print("叠加次数 N", N)

def te_func(weights, returns, index_returns):    # tracking error
    error_sum = 0
    for index, row in returns.iterrows():
        c_sum = weights @ np.array(row).T
```

```
        error_sum += (c_sum - index_returns[index]) ** 2
    return error_sum / len(returns)

def er_func(weights, returns, index_returns):    # Excess return
    excess_returns = 0
    for index, row in returns.iterrows():
        w_sum = weights @ np.array(row).T
        excess_returns += w_sum - index_returns[index]
    return excess_returns / len(returns)

def min_fun(weights, returns, index_returns):
    trend_error = te_func(weights, returns, index_returns)
    excess_returns = er_func(weights, returns, index_returns)
    return trend_error - (1 - CUSTOM_T) * excess_returns

def constraint1(weights, esg):    # >=CUSTOM_ESG
    esg_t = np.array(esg).T
    weight_esg = weights @ esg_t
    return weight_esg - CUSTOM_ESG

def constraint2(weights):
    return np.sum(weights) - 1

def constraint3(weights):    # <=K
    zero_count = [ 0 if x != 0 else 1 for x in weights]
    return K - np.sum(zero_count)

if __name__ == '__main__':
    df_esg = pd.read_excel(f'{INPUT_DIR}/esg.xlsx', sheet_name="ESG_富时罗素")    # ESG
data loading
    # df_esg['Date'] = df_esg['Date'].apply(lambda x:str(x)[0:7])
    df_esg.set_index('Date', drop=True, inplace=True)

    df_index = pd.read_csv(f'{INPUT_DIR}/index.csv')    # index benchmark return

    df_returns = pd.read_excel(f'{INPUT_DIR}/收益率.xlsx', sheet_name='Sheet2')    # stock
return
    df_returns['Date'] = df_returns['Date'].apply(lambda x:str(x)[0:10])
    df_returns = df_returns.merge(right=df_index[['日期', 'rate']], left_on="Date", right_on="日
期", sort=True)
    df_returns = df_returns.loc[df_returns['日期'] > '2019-08-31']
    # df_returns['month'] = df_returns['日期'].apply(lambda x:x[0:7])

    # returns_group = df_returns.groupby(by='month')
```

```
result = []
for i in range(42, len(df_returns) + 1, 1):
    m_returns = df_returns.iloc[i - 42:i]
    N = 0
    m_returns.set_index('Date', drop=True, inplace=True)
    # remove empty rows
    m_returns = m_returns.dropna(axis='columns')
    returns_index = m_returns['rate']    # index monthly return
    m_returns.drop(columns=['日期', 'rate'], inplace=True)    # stock monthly returns

    c_date = m_returns.iloc[-1].name

    esg_month = df_esg.loc[c_date:].iloc[0]
    esg_month = esg_month.dropna()
    # intersection between esg and returns
    intersec_col = list(set(m_returns.columns) & set(esg_month.index))
    intersec_col.sort()
    print(c_date)
    weights_dic = {'date':c_date}
    weights = get_weights(m_returns[intersec_col], returns_index,
esg_month.loc[intersec_col])
    weights_dic.update(weights)
    result.append(weights_dic)
df = pd.DataFrame(result)
df.to_csv(OUTPUT_FILE, index=False)
print(df)
```