ISSN 2959-6130

Stock Selection Analysis of Innovative Multifactor Model in New Energy Sector

Jiaxuan Shao

International education,Hebei University of Economics and Business,Shijiazhuang, Hebei, 050000, China
Email: sjx15100117616@outlook.com

Abstract:

As of 2024, China's new energy sector has made significant progress, ranking first in the world in terms of installed photovoltaic (PV) and wind power capacity. The proportion of new energy in the energy mix has increased significantly, driving China's energy transition and sustainable development. Meanwhile, with the rapid development of Internet technology, quantitative investment strategies based on artificial intelligence technology are expected to show great growth potential in the future financial sector. In this context, this paper will construct a multi-factor model with the help of machine learning, combine value analysis and technical analysis, select candidate factors, develop exclusive factors for the characteristics of the new energy industry, test the validity of the factors, construct a multi-factor stock picking strategy and carry out backtesting analysis. The results demonstrate that the strategy can effectively manage risks and achieve excess returns, which is theoretically and practically instructive for the construction of investment portfolios. The results provide an effective investment strategy reference for financial institutions such as banks, brokerage firms and fund companies.

Keywords: new energy, quantitative investment, multifactor modelling, IC value approach

1. Introduction

As of 2024, China's new energy sector is growing rapidly, driven by multiple factors such as environmental protection, energy security, technological advances, policy support, market demand, international cooperation, and the Sustainable Development Goals (SDGs). As energy demand continues to grow as a result of rapid economic development and urbanisation, new energy sources not only meet the increasing demand for energy, but also promote economic growth and job creation. Against this backdrop, .the performance of the new energy sector in the stock market has garnered widespread attention. How to effectively conduct stock selection in this emerging sector has become a core concern for investors and researchers. The traditional Capital Asset Pricing Model (CAPM) and single-factor model have certain limitations in practical application, especially when dealing with the new energy sector, which has high growth and volatility. This study expands the research scope of stock picking models through the construction of innovative multi-factor models, provides a theoretical framework and empirical analysis method for new energy sector, and provides new perspectives and new ideas for academic research in related fields.

2. Literature Review

In foreign countries, for the multi-factor model related research started earlier, Markowitz (1952) first proposed portfolio mean-variance analysis model, successfully reduced portfolio risk, laid a solid foundation for modern asset allocation theory [1]. Ross (1976) developed a more comprehensive Arbitrage Pricing Theory (APT), which suggests that asset prices are influenced by multiple factors rather than being explained by a single systematic risk [2]. The formulation of this theory has facilitated in-depth research on multi-factor models. In particular, Fama and French (1993) observed that a single beta coefficient could not adequately explain the differences in stock returns, so they added two new factors - book-to-market ratio and market value - to construct a more accurate three-factor model [2]. Fama and French (2015) extend their model again by adding two factors, profitability and investment style, to form a more comprehensive five-factor model. Carhart (1997) added momentum factor to the three-factor model and created a four-factor model including inertia effect [3].

Domestic research on multi-factor modelling started later than international research, but has gained momentum in recent years. Liu Yi (2012) conducted an empirical study based on data from 2000 to 2012 and selected four key factors, namely, growth, valuation, quality, and momentum, and identified eight outperformers from them to construct the optimal combination of factors. This portfolio shows good performance under different market conditions. In terms of the development of new factors, the research on macro factors, fundamental factors and statistical factors has matured, and in recent years, more and more researchers have begun to construct statistical factors on their own. For example, Huang Yanjing (2017) developed an investor sentiment factor for the A-share market, selecting four key variables including the market turnover rate and the price-earnings ratio, which were incorporated into a four-factor model after excluding the impact of macroeconomic cycles, showing better performance than a single sentiment factor such as the number of newly established funds.

Scholar Qian Bozhang (2007) mentioned that in order to cope with the challenge of high oil prices and to meet the arrival of the post-petroleum era, it is necessary to accelerate the development of alternative energy sources in order to realise the sustainable development of the human society [4] . Zhu Xiping and Chen Ying (2010) discussed in detail how to promote the development of new energy industry through the reform of financial policies and investment and financing mechanisms. Wang Shuping and Yan Xiaofeng (2010) analysed in depth the practices of developed countries in the formulation of new energy policies and strategic planning, and extracted the international experiences in the light of China's specific national conditions, which provided a series of specific references and suggestions for optimising China's new energy industry policies[5]. Jin (2011) highlighted the importance of international cooperation in the development of new energy industries, pointing out that new energy is an important means to cope with climate change and to move towards a low-carbon economy [6]. He Mang (2019) systematically introduced the current situation of China's new energy industry development and the existing technical, resource and information barriers, and put forward the suggestion that the development of new energy must vigorously cultivate talents of new energy technology and raise the people's awareness of new energy[7] .

3. Experimental setup

3.1 *Data collection*

This paper takes 000941.XSHG new energy stock as the research object, and collects the annual reports of these 50 listed companies in the decade of 2014-2023 from Sina Finance website and other media websites.

3.2 *Data pre-processing*

The annual reports from the Sina Finance website are in PDF format and require formatting. This step is performed on the Xindao cloud platform. Check whether the document can be txt converted by running the code. As shown in Figure 1:

```
import os
from pdfminer.pdfparser import PDFParser
from pdfminer.pdfdocument import PDFDocument
from pdfminer.pdfinterp import PDFResourceManager, PDFPageInterpreter
from pdfminer.converter import PDFPageAggregator
from pdfminer.layout import LTTextBoxHorizontal, LAParams
from pdfminer.pdfdocument import PDFTextExtractionNotAllowed
from pdfminer.pdfpage import PDFPage
import logging
import re


logging.propagate = False
logging.getLogger().setLevel(logging.ERROR)

pdf_folder = "C:/Users/DELL/Desktop/test/pdf"
txt_folder = "C:/Users/DELL/Desktop/test/txt"

if not os.path.exists(txt_folder):
    os.makedirs(txt_folder)

pdf_files = []
for file_name in os.listdir(pdf_folder):
    file_ =os.path.join(pdf_folder, file_name)
    file_name_list = os.listdir(file_)
    for file_n in file_name_list:
        file_path = os.path.join(file_, file_n)
        pdf_files.append(file_path)
```

Figure 1 Annual report format conversion

3.3 *Data analysis and mining*

jieba library is a practical Python third-party Chinese lexical library, through the Python programming operations, you can lexically slice Chinese text, to facilitate text analysis of the processed text. Regarding the algorithm, the jieba library performs efficient word map scanning based on the prefix lexicon, and generates a directed acyclic graph (DAG) consisting of all possible word formation cases of Chinese characters in a sentence; it adopts dynamic programming to find the path of maximum probability, and finds out the maximum combination of cuts based on the frequency of the word; for the unregistered words, it adopts the Viterbi algorithm based on the HMM model of Chinese character's ability to form words. Operationally, it is not necessary to understand the principle of the algorithm when calling the jieba library, but only need to know that jieba can decompose the article into an independent word

Because of the completion of the word segmentation work, the word statistics list will appear in many "I", "the" and so on the frequency of high but no significant role in the semantic analysis of the word, this paper will be called "deactivated words This paper refers to them as "deactivated words". In addition, some words in jieba's own thesaurus are cut off, due to the financial vocabulary itself has special characteristics, so this paper through the construction of a more targeted financial thesaurus to split the words.

In the use of jieba thesaurus for segmentation after the completion of the thesaurus has been segmented word frequency statistics to derive the current text, excluding the deactivation of the number of occurrences of participles, word frequency statistics. The first step of the text segmentation operation: retrieve the segmented and deactivated thesaurus for iieba segmentation. Import the relevant Python libraries on the Xindao cloud platform, open the document that needs to be counted for word frequency, and then call the special financial thesaurus to carry out the word division operation on the text, and at the same time establish a counting framework, and finally call the deactivated word thesaurus.

The second step of the word frequency statistics operation operation: Secondary processing and counting of participle words. Firstly, the counting loop is established, secondly, the deactivated words/single-word participles are eliminated, next, the counting loop is carried out, the results are sorted, the high-frequency words are extracted, and finally, the results are exported.

3.4 *Creating new factors*

According to the word frequency statistics, the keywords closely related to the development of new energy are screened, such as 'photovoltaic', 'battery', 'new energy' and 'research and development', etc., and these four words are constructed into a special factor named 'energy', and the word frequency of each enterprise in each decade is counted. This is shown in the Figure 2 below:

```
import multiprocessing as mp
import os
from loguru import logger
import jieba
import pandas as pd
import pdfplumber


def opne_read_file(file_path):
 text_all = ""
with pdfplumber.open(file_path) as pdf:
 for page in pdf.pages:
 text_all += page.extract_text()
name_2 = file_path.split('\\')[-2]
name = file_path.split('\\')[-1].split('.')[0]

data = pd.DataFrame({"data": list(jieba.cut(text_all))})

data = data[data['data'].map(lambda x: x in [ 'photovoltaic', 'battery', 'new energy' and 'research and development']
0].sum() / 103630
return {"name": name_2, "time": name, "data": data}


if __name__ == '__main__':
 path = "New Energy Stocks Annual Report"
 file_file = os.listdir(path)
  file_path_ = []
 for file_name in file_file:
 file_ = os.path.join(path, file_name)
```

Figure 2 Generation of new factors

4. Factor analysis based on the Alphalens framework

4.1 *Selection of Sample Data*

In this paper, we empirically test the multi-factor model based on the new energy constituent stocks updated after January 2014, and the sample stocks selected are the constituent stocks of 000941.XSHG new energy in the MiBasket Quantitative Collaboration Platform. In order to objectively evaluate the return of the constructed multi-factor stock picking model, we adopt CSI 300 index as the market benchmark. There are two main reasons for choosing CSI 300 index as the evaluation benchmark: firstly, CSI 300 index can comprehensively reflect the stock price changes of the representative stocks with good liquidity and large scale in the market, which accurately reflects the overall trend of the A-share market, and therefore can be used as an effective standard to measure the investment performance. Secondly, the industry distribution of CSI 300 Index is basically consistent with the proportion of industry distribution of the whole market, which is highly representative. Therefore, CSI 300 Index is not only an effective benchmark for measuring market performance, but also provides a reliable reference for investors' investment decisions.

The sample data period selected for this paper is from 1 January 2014 to 31 December 2023, and the data selected is daily frequency data for a total of 10 years (i.e., 2,430 trading days), and the source of this sample data is the MiBasket Quantitative Collaboration Platform, and some of the data is shown in the Figure 3 below:

	order_book_id	date	operating_revenue	total_assets	cash_equivalent	bill_accts_receivable	pe_ratio_ttm	market_cap_2	market_cap_3	ev_ttm
0	600732.XSHG	2014-01-02	1.528099e+07	9.581664e+08	3.873507e+07	NaN	-1671.610770	2.339047e+09	2.339047e+09	3.057977e+09
1	600732.XSHG	2014-01-03	1.528099e+07	9.581664e+08	3.873507e+07	NaN	-1671.610770	2.339047e+09	2.339047e+09	3.057977e+09
2	600732.XSHG	2014-01-06	1.528099e+07	9.581664e+08	3.873507e+07	NaN	-1671.610770	2.339047e+09	2.339047e+09	3.057977e+09
3	600732.XSHG	2014-01-07	1.528099e+07	9.581664e+08	3.873507e+07	NaN	-1671.610770	2.339047e+09	2.339047e+09	3.057977e+09
4	600732.XSHG	2014-01-08	1.528099e+07	9.581664e+08	3.873507e+07	NaN	-1671.610770	2.339047e+09	2.339047e+09	3.057977e+09
...	...	...	...	...	...	...	...	...	...	...
121695	688707.XSHG	2023-12-25	5.280979e+09	8.749837e+09	1.415604e+09	1.273550e+09	28.016374	NaN	8.508778e+09	1.412170e+10
121696	688707.XSHG	2023-12-26	5.280979e+09	8.749837e+09	1.415604e+09	1.273550e+09	27.433004	NaN	8.331604e+09	1.394452e+10
121697	688707.XSHG	2023-12-27	5.280979e+09	8.749837e+09	1.415604e+09	1.273550e+09	27.243408	NaN	8.274022e+09	1.388694e+10
121698	688707.XSHG	2023-12-28	5.280979e+09	8.749837e+09	1.415604e+09	1.273550e+09	29.358127	NaN	8.916278e+09	1.452920e+10
121699	688707.XSHG	2023-12-29	5.280979e+09	8.749837e+09	1.415604e+09	1.273550e+09	29.897745	NaN	9.080164e+09	1.469308e+10

121700 rows × 24 columns

Figure 3 Specific values of the factors

4.2 *Candidate factor selection and processing*

4.2.1 *Selection of candidate factors*

The candidate factors in the paper are divided into technical and financial indicators. Among them, in technical indicators, this paper constructs new factors by text data specially. The selection of financial indicators, on the other hand, mainly relies on economic logic and market experience. Among the nine candidate factors, they are divided into one technical indicator and eight financial indicators, and the details are shown in Table 1.

Table 1 Factor selection based on Alphalens framework factor analysis

<table>
<tr><th colspan="2">factor type</th><th>field</th><th>Factor name</th></tr>
<tr><td>Technical indicators</td><td>Text indicators</td><td>energy</td><td>Energy indicators</td></tr>
<tr><td rowspan="8">Financial Indicators</td><td rowspan="2">Snapshot data</td><td>operating_revenue</td><td>revenues</td></tr>
<tr><td>total_assets</td><td>total assets</td></tr>
<tr><td rowspan="2">Underlying financial data</td><td>cash_equivalent</td><td>money funds</td></tr>
<tr><td>bill_accts_receivable</td><td>Notes and accounts receivable</td></tr>
<tr><td rowspan="4">Derivative financial indicators</td><td>pe_ratio_ttm</td><td>PE ratio</td></tr>
<tr><td>market_cap_2</td><td>Total market value of shares outstanding</td></tr>
<tr><td>market_cap_3</td><td>total market value</td></tr>
<tr><td>ev ttm</td><td>enterprise value</td></tr>
</table>

In addition to the technical indicators specially constructed in this paper,this paper selects the first four financial indicators listed in the table, with the selection criteria grounded in three key principles: first, the representativeness of the financial indicators. The financial indicators selected in this paper can intuitively reflect the company's operating conditions and financial status, and can quickly understand the company's profitability, solvency, operational efficiency and growth potential and other key information. These indicators help identify high-quality stocks with investment value, so as to make more informed investment decisions; second, the stability and predictability of financial indicators. By analysing historical financial data, some financial indicators with stable performance can be identified, and these indicators can predict the future development trend of the company to a certain extent. In the process of stock selection, this helps to reduce the risk and improve the investment return; third, the accessibility of financial indicators. The selected factor data belongs to the public information of the A-share market, and the frequency of information is daily, monthly and quarterly, etc.. The public data is easy to obtain and can be used as the basis for investors' investment decisions, reducing information asymmetry. This can make it easier to use the multi-factor stock selection strategy for stock selection.

4.2.2 *Processing of candidate factors*

The initial step involves sorting the data chronologically. In the previous step, the candidate factor values we obtained were sorted by the default index (i.e., stock code). However, in order to satisfy the subsequent data processing requirements, it is necessary to convert this data structure into a DataFrame that is sorted in chronological order. after performing the corresponding conversion operation, the result obtained is shown in Figure 4:

		operating_revenue	total_assets	cash_equivalent	bill_accts_receivable	energy
date	**order_book_id**					
2014-01-02	**603659.XSHG**	NaN	NaN	NaN	NaN	0.005761
	603799.XSHG	NaN	NaN	NaN	NaN	0.000743
	300073.XSHE	4.558349e+08	9.775052e+08	9.864249e+07	2.573426e+08	0.001602
	001289.XSHE	NaN	NaN	NaN	NaN	NaN
	300014.XSHE	7.014151e+08	1.184675e+09	1.898689e+08	3.422494e+08	0.001245

Figure 4 Factor values sorted by time

Step 2, factor standardisation

First, the data were processed for extreme values by replacing the extreme values (i.e., minimum and maximum values) of each candidate factor with the corresponding quantile values (i.e., 2% and 98% quantiles). For values less than the minimum quartile, they are replaced with the minimum quartile; for values greater than the maximum quartile, we replace them with the maximum quartile. After this treatment, the new variables are obtained. Secondly, standardisation is carried out, i.e. Z-Score standardisation method is used. This method is based on the mean and standard deviation of the original data for standardisation. After this treatment, the data will conform to the standard normal distribution, i.e., its mean becomes 0 and standard deviation becomes 1. Such a treatment helps in the subsequent data analysis and modelling work.

The transformation function is:

$$x_{new} = \frac{x - \mu}{\sigma}$$

where μ is the mean (mean) of the sample data and σ is the standard deviation (std) of the sample data.

The normalisation process serves to, firstly, improve the speed of convergence of the model (speeding up the solution of gradient descent); secondly, to improve the accuracy of the model (eliminating the effects of magnitude and magnitude); and thirdly, to simplify the computation.

View the results of the Energy Factor process as shown Figure 5 below:

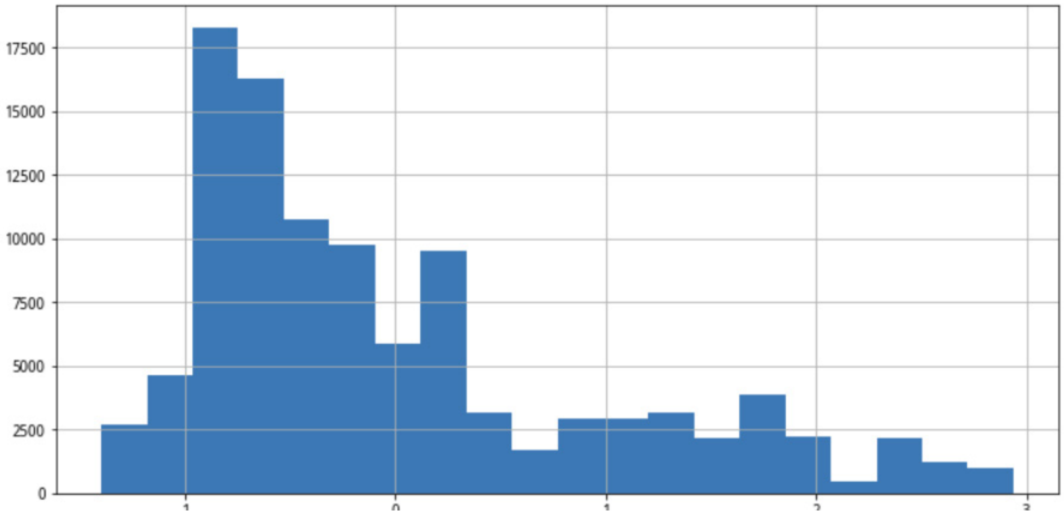

Figure 5 Standardised histogram of energy factors

4.3 *One-way validity tests*

4.3.1 *Factor analysis with Alphalens*

For formatting input data, Alphalens provides a handy function that is primarily used to convert factor and price data into the specific format required by the chart function. During the formatting process, the function reports in detail how much data has been discarded. In fact, some of the factor data is discarded by for a variety of reasons, such as missing data (NaN values), lack of sufficient price data to compute forward returns, inability to group, etc. In order to effectively manage the data discard situation, this paper sets an upper limit to control how much data can be discarded at most through the max_loss parameter, and the processing is shown in Figure 6:

格式化输入数据 Alphalens包含一个很方便的函数，用来把你的因子和价格数据转换成图表函数所需要的格式。

factor_data = alphalens.utils.get_clean_factor_and_forward_returns(predictive_factor, pricing, quantiles=5, bins=None, groupby=ticker_sector, groupby_labels=sector_names)

这个函数会告诉用户格式化过程中有多少数据被丢弃。部分因子数据会出于若干种原因被丢弃，比如数据为空（NaN）、没有足够的价格数据来计算前向收益、没法分组等等。我们也可以通过max_loss参数来控制最多丢弃多少数据。

```
facs_data_analysis_operating_revenue = utils.get_clean_factor_and_forward_returns(series_facs_datas_operating_revenue,price)
facs_data_analysis_total_assets = utils.get_clean_factor_and_forward_returns(series_facs_datas_total_assets,price)
facs_data_analysis_cash_equivalent = utils.get_clean_factor_and_forward_returns(series_facs_datas_cash_equivalent,price)
facs_data_analysis_bill_accts_receivable = utils.get_clean_factor_and_forward_returns(series_facs_datas_bill_accts_receivable,price)
facs_data_analysis_energy = utils.get_clean_factor_and_forward_returns(series_facs_datas_energy,price)
```

```
Dropped 28.3% entries from factor data: 28.3% in forward returns computation and 0.0% in binning phase (set max_loss=0 to see potentiall
y suppressed Exceptions).
max_loss is 35.0%, not exceeded: OK!
Dropped 28.3% entries from factor data: 28.3% in forward returns computation and 0.0% in binning phase (set max_loss=0 to see potentiall
y suppressed Exceptions).
max_loss is 35.0%, not exceeded: OK!
Dropped 28.4% entries from factor data: 28.4% in forward returns computation and 0.0% in binning phase (set max_loss=0 to see potentiall
y suppressed Exceptions).
max_loss is 35.0%, not exceeded: OK!
Dropped 28.4% entries from factor data: 28.4% in forward returns computation and 0.0% in binning phase (set max_loss=0 to see potentiall
y suppressed Exceptions).
max_loss is 35.0%, not exceeded: OK!
Dropped 33.8% entries from factor data: 33.8% in forward returns computation and 0.0% in binning phase (set max_loss=0 to see potentiall
y suppressed Exceptions).
max_loss is 35.0%, not exceeded: OK!
```

Figure 6 Alphalens processing

4.3.2 *Calculation of IC values*

After calculating the IC value of the candidate factor, it is compared with the specified value and assign a score according to the following calculation rules: when IC-mean>0.01, it is scored as 0.5, otherwise it is scored as 0; when ICstd>0.01, it is scored as 1.0, otherwise it is scored as 0; when (IC>0.02)>0.48, it is scored as 0.5, otherwise it is scored as 0; when IR& gt;0.01, score 1.0, otherwise score 0; factor return >0.000, score 3, otherwise score 0. The total score is 6 points. The results of the calculations were further ranked as shown in Figure 7 below.

	IC mean	IC std	IC>0.02	IR	Factor Yield	Factor total score
FactorCriterion	0.010000	0.100000	0.480000	0.010000	0.000000	0.0
CreditPoint	0.500000	1.000000	0.500000	1.000000	3.000000	0.0
energy	0.007771	0.211789	0.473597	0.036694	0.000444	5.0
operating_revenue	0.007374	0.211759	0.479373	0.034823	-0.002333	2.0
total_assets	0.007698	0.219167	0.474422	0.035122	-0.002493	2.0
cash_equivalent	0.008529	0.209137	0.481848	0.040783	-0.001941	2.5
bill_accts_receivable	0.010840	0.219863	0.479785	0.049304	-0.002040	2.5

Figure 7 Calculation of IC value results

The test shows that only the factor return of innovation is positive and it has a score of 5.0, which can indicate the validity of its indicator.

4.4 *Building a multi-factor composite scoring model and stock selection*

The construction of a multi-factor stock picking portfolio involves four key steps, the first step is to select candidate factors, the second step is to test the validity of stock picking indexes, the third step is to assign weights to the valid factors and build a composite scoring model, and the fourth step is to conduct a time-timed backtesting of the stocks, calculate the return and conduct validity test and result analysis.

In this step, this paper uses two stock pools to construct portfolios for strategy backtesting. The first stock pool is constructed by selecting the factor with the highest total factor score (5), the energy factor, and performing stock selection to obtain the top 10% of stocks in its size as the stock pool. The second also uses this highest scoring factor for stock selection as well, with the difference that it expands its range to 20% of stocks for the stock pool.

4.5 *Construction of stock pools*

Select the highest factor score (5.0) of the energy factor selection for stock selection, to obtain 000941.XSHG new energy stock constituents on 1 January 2014energy factor indicator size of the top 16 stocks as a stock pool, the selected 16 stocks are: 300037.XSHE new yupang, 300438. XSHE penghui energy, 603185 XSHG Hongyuan Green Energy, 688778.XSHG Xiamen Tungsten New Energy, 688707.XSHG Zhenhua New Material, 688779.XSHG Longyuan Lithium Science and Technology, 001301. XSHE Shangtai Science and Technology, 600905.XSHG Three Gorges Energy, 688599.XSHG Trina Solar, 688223. XSHG JinkoSolar. 300274. XSHE Sunny Power, 603659. XSHG PuTaiLai, 300750.XSHE Ningde Times, 600438. XSHG Tongwei, 002129.XSHETCL Zhonghuan, 688005. XSHG Rongbai Technology.

4.6 *Selection of timing strategies and backtesting results*

In conjunction with the above stock pool, the following strategy is again adopted for timing backtesting, and the selection of indicators for the strategy is described in detail below. The above 16 stocks are rotated on a regular basis with a full position cycle of 20 trading days, i.e. the pool will be updated once every 20 trading days.

The backtest results for this strategy are as shown in Figure 8 belolw:

Figure 8 Backtesting results

The paper selects the CSI 300 index as the benchmark, and in the backtest of this strategy, the benchmark return is 47.256% and the benchmark annualised return is 4.088%. In contrast, the portfolio constructed in this paper gets better investment backtest results. Specifically, the backtest return of this paper's strategy is 582.740%, and the backtest annualised return is 22.004%. The annual win rate is 0.6000, the alpha is 0.1795, the beta is 1.0603, the Sharpe ratio is 0.5953, and the maximum retracement is 58.723%. It can be seen that this investment strategy has achieved very high returns during the backtesting period, but it also shows a high level of risk. Where the larger maximum retracement may mean that the strategy may suffer larger losses in high volatility or unfavourable market conditions. Therefore, these factors need to be carefully considered in practical application and decisions need to be made in conjunction with one's own investment objectives and risk tolerance.

5. Conclusion and outlook

After a series of rigorous factor selection, validation and redundant factor removal processes, we have successfully constructed a multi-factor stock selection model based on five core factors. The model's performance has remained stable over most of the long-term backtesting period, and its return even outperformed the CSI 300 index, showing that its portfolio outperformed the market as a whole. The stock selection model constructed in this study is somehow able to reduce the irrational decisions made by investors due to emotions or biases, thus enabling investors to secure more robust excess returns. At the same time, the selection of factors is mainly based on their actual performance in the market, which to a certain extent weakens the influence of decision-making differences between different investors.

However, our study also found that the current multi-factor stock selection model has some limitations, especially it fails to fully take into account the dynamic variability of the factors. In fact, the validity of factors changes over time. In certain short-term periods, some originally valid factors may become ineffective, while some factors that were previously considered invalid may show better performance in the current market environment. Therefore, it is particularly important for institutions such as banks and wealth management subsidiaries to make continuous dynamic adjustments and optimisation of factors.

Financial institutions should actively respond to market changes, constantly iteratively update their quantitative strategy factors, and at the same time conduct in-depth research on the unique asset characteristics of the new energy industry. Based on these studies, they should develop factors that are truly suitable for the new energy sector, and explore the inner rules of the effectiveness of these factors. In addition, we need to continue to explore and improve multi-dimensional quantitative investment strategies that are compatible with the development of the new energy industry, and scientifically predict the effective duration of the strategy models, so as to inject strong momentum into the vigorous development of the energy industry capital market.

In the ever-changing new energy sector,we should comprehensively assess the market performance and risk characteristics of various sub-sectors, by means of sophisticated quantitative models. At the same time, we should strengthen the use of big data analysis to develop and rigorously verify various trading strategies. With the help of advanced artificial intelligence technology, the annual reports and other important announcements of new energy listed companies are analysed in depth, key information is extracted accurately, and scientific and effective investment strategies are constructed accordingly. In addition, it is necessary to regularly evaluate the performance of the quantitative stock picking strategies employed to comprehensively ensure that the strategies are able to achieve the desired investment objectives.

To promote inter-industry collaboration, communication and co-operation between financial institutions and the new energy industry should be strengthened to achieve resource sharing and complementary advantages, and to jointly promote the development of the new energy sector. By constructing a systematic and scientific multi-factor quantitative stock selection strategy for the new energy industry, and backtesting and optimising it on a regular basis, in order to better adapt to the constant changes in the market and the continuous development of the industry. At the same time, we provide investors with more professional and caring asset management services to help them move forward steadily in the complex and volatile market environment.

Strengthen investor cultivation. Through in-depth education, we should guide investors to deeply understand the intrinsic connection between the long-term growth potential of the new energy industry and short-term market volatility, so that they can make more informed investment decisions.

6. Acknowledgments

I would like to express my special thanks to Mr Sun Wei for his technical support during the research process and for the revisions he provided during the writing of the thesis.

References

[1] H. Markowitz. Portfolio selection [J]. The journal of finance, 1952, 7(1): 77-91.

[2] Li Heting. Research on quantitative stock selection of A-share new energy vehicle sector based on multi-factor model[D]. East China Normal University,2022.

[3] Fama, E. F and K. R. French (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33(1), 3-56.

[4] Carhart, M.M. (1997).On persistence in mutual fund performance. Journal of Finance 52(1), 57-82.

[5] Qian Bozhang. New energy - the inevitable choice in the post-oil era [M]. Beijing:Chemical Industry Press,2007.

[6] WANG Shuping,YAN Xiaofeng. New energy policies of major developed countries and their implications for China[J]. Proceedings of the Fourth International Forum on China’s Energy Strategy, 2010.

[7] Jin. China’s new energy development report [M]. Huazhong University of Science and Technology Press, 2011(07).