

# *Research on the Construction and Tracking Method of Climate Risk Index Based on News Reports*

Simin Yang\*, Zile Chen, Bowei Zhan, Jing Guan, Yihan Chai

*International Business College, South China Normal University, Foshan, China*

*\*Corresponding author: anicaxd@outlook.com*

**Keywords:** Climate risk, Financial risk management, Index Construction and Tracking

**Abstract:** Climate change has become one of the important factors affecting the structural changes of the financial system, and the risks related to climate change have also become one of the sources of financial risks. Based on the perspective of news reports, this paper constructs two climate risk indexes to track and study the risks brought by climate change and the impact of climate risks on the financial sector, so as to lay a solid foundation for preventing financial systemic risks.

## 1. Introduction

In recent years, extreme weather has been increasing. Climate trajectory and climate change have brought great economic losses and uncertainties, increased market risks and liquidity risks, and also affected the implementation space and transmission channels of monetary policy. In this context, climate finance related to climate change and financial development is becoming an important academic frontier of financial research.

For the central bank, climate change undoubtedly brings challenges to monetary policy and financial stability, and is an important issue that the monetary authorities of major economies in the world have to face; For investors, they also have a huge demand in hedging climate risks, so it is necessary to discuss how to establish a climate risk index that can be effectively tracked.

## 2. Index Construction

We believe that we should first build a climate related time series to capture the news about climate risks, so as to build an appropriate hedging target. According to the observation, newspapers usually report events related to climate risk change, and related topics include extreme weather events, geophysical changes, regulatory discussions, technological progress in alternative fuel transmission, and the price of fossil fuels. Many investors obtain information about climate from this. Therefore, the method we use in this paper is to extract a climate news time series from the text analysis of news sources. Various events reported in newspapers may carry potentially relevant climate information. We refer to the construction method of the climate risk index in the article *Heading Climate Change News*, and build two indexes according to China's climate and data conditions, so as to measure the degree of discussing climate change in the news media.

This index is called the climate risk index, which also refers to the index of climate change

reports below. It is an index used to track negative climate change based on Climate News reports and under the assumption that increased climate change reports mean increased climate risk. This index can effectively help investors establish climate risk hedge positions by buying stocks with low carbon footprints when the climate risk index rises and selling stocks with high carbon footprints, or buying stocks with high carbon footprints when the climate risk index falls and selling stocks with low carbon footprints to build portfolios.

### 3. Calculation Method

The climate risk index is calculated as follows:

#### 3.1. Screening climate change keywords with Chinese characteristics

We first screened the top 160 key words of climate change in these white papers based on 12 authoritative climate change white papers in Europe, America and the world, and built a preliminary key thesaurus, and then added a Chinese specific keyword pool suitable for China's national conditions and development status. Finally, we use these words to compare the climate news of more than 50 news websites such as China Daily and China News Network in the past year, select the keywords with the top 80 frequencies in the preliminary key vocabulary, and construct the final key vocabulary.

#### 3.2. Collect sufficient and non-repetitive news information

At 24:00 p.m. every day, we crawl all the news from more than 50 news websites, such as China Daily and China news. Each news constitutes an independent sub document. We use machines to identify and delete duplicate news sub documents, and simply integrate the filtered sub document library into one document as the general document of the day.

#### 3.3. Calculate the daily TF-IDF value of keywords that appear on that day

The calculation formula is:

Term Frequency (TF) = the number of occurrences of a keyword in the total document / the total number of words in the total document

Inverse Document Frequency (IDF) =  $\log(\text{total number of sub documents} / \text{number of sub documents containing the keyword} + 1)$

TF-IDF = Term Frequency (TF) \* Inverse Document Frequency (IDF)

It can be seen from the formula that TF-IDF is directly proportional to the number of occurrences of a word in the total document and inversely proportional to the number of occurrences of the word in the whole different sub documents, which shows that the index not only effectively considers the frequency of occurrence of each keyword, but also avoids the multiple occurrence of some single keywords that are not important enough to affect the whole. The calculation method of TF is: calculate the number of occurrences of keywords in the total document of the day / the total number of words in the total document of the day; IDF is calculated as:  $\log(\text{total number of sub documents in the current day} / \text{number of sub documents containing the keyword in the current day} + 1)$ .

#### 3.4. Calculate the daily TF-IDF value of the key vocabulary

The TF calculation method of keywords is to calculate the frequency of each keyword in the key

vocabulary. The IDF calculation method of keywords shall be the same as that of keywords appearing on the same day to ensure the weight of file frequency.

### 3.5. Calculate cosine similarity

We calculate the TF-IDF value of each keyword in the daily total document and make it into a vector of 80 dimensions, which is recorded as  $F1 = (k1a1, k2a2, k3a3, \dots, k80a80)$ , where  $k_n$  is the word frequency of the  $n$ th keyword and  $a_n$  is the inverse document frequency of the  $n$ th keyword. Take the daily TF-IDF value of the key vocabulary as a vector of 80 dimensions, record it as  $F2 = (a1, a2, a3, \dots, a80)$ , and calculate the cosine  $\cos(F1, F2)$  between the two vectors.

### 3.6. Build climate risk index

We magnify  $\cos(F1, F2)$  by 10000 times as a climate risk index to track daily climate risk changes.

### 3.7. Construction of sub index

We select E(environment) in ESG, which is an important indicator in the environment, including power generation, power consumption, carbon emissions, etc. the same as the calculation steps of the climate risk index. We select different news databases and key thesaurus, and use the same method to build a sub index of the climate risk index to track more detailed climate risk changes.

## 4. Index Construction Again

In F.E. Robert, G. Stefano, T.K. Bryan, L. Heebum and s Johannes' article *Heading Climate Change News* puts forward a more intuitive negative news index of climate change. In the process of index construction, the article directly distinguishes between positive news and negative news. Therefore, we have built a second news-based climate index, which focuses on negative news about climate change. This index applies emotional analysis to articles to measure the intensity of negative Climate News in a certain day.

We searched the economic news database of Infobank for the relevant phrase Climate Change, and limited our search results to news media that do not include social media. Based on the choice of these terms and content sources, we have achieved some results on the economic news database of Infobank database, summarizing the total number of articles including climate change news and the proportion of articles containing positive and negative climate change news. It also provides an index for further emotional subcategories (such as fear, happiness, anger, etc.). Then we divide the number of "negative emotion" news obtained from the economic news database of Infobank database by the number of news with climate change as the keyword on that day, and multiply this result by 10000 as the index of climate change negative news.

## 5. Summary

Because Chinese data websites generally lack accurate emotion recognition function, we use the test method in the article *Heading Climate Change News* to compare the index of climate change reports in the United States from 2008 to 2018 with the broken line chart of the index of climate change negative news. The construction method and principle of these two indexes are basically consistent with this article. It can be found from the picture that the index of climate change reports is roughly the same as that of climate change negative news, except for the period from 2008 to

2009, and the ch sample size of the news database that constructs the index of climate change negative news during this period is small. Because the trend of the index of climate change reports is close to that of the negative news of climate change, we have sufficient reasons to believe that the increase of climate change reports means the rise of climate risk. Even in some major climate events, the index of climate change reports is better than that of the actual negative news of climate change.

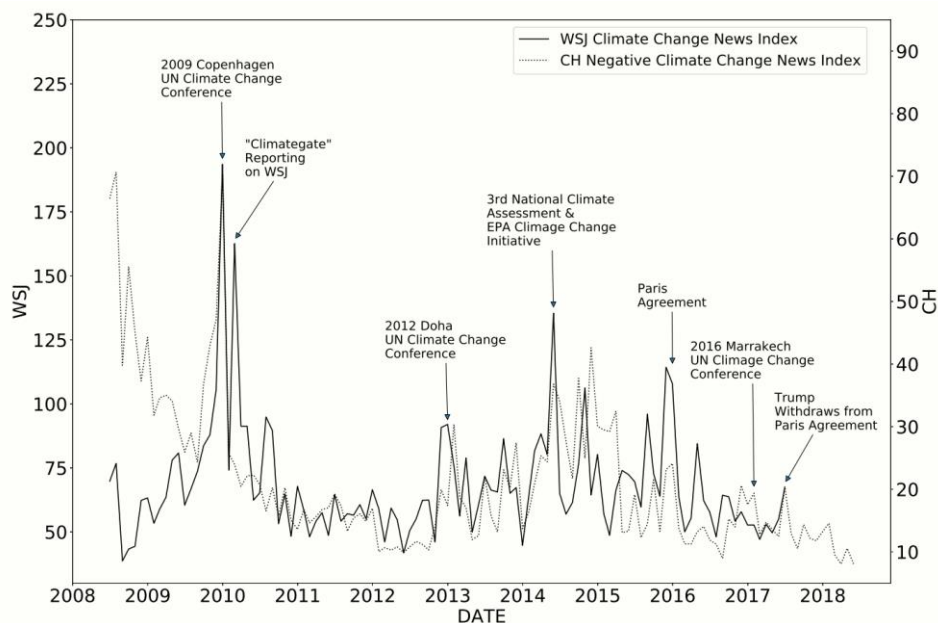


Figure 1: Comparison between climate change negative news index and climate change report index

According to our analysis, it is possible that the actual climate change negative news index will selectively abandon some neutral climate reports or both positive and negative climate reports when it is constructed, resulting in its sensitivity being inferior to the climate change report index that has not been artificially screened. Therefore, we choose Climate News reports to directly track climate risk change.

## References

- [1] Zhang Wenjuan, Yang Cuo. *Climate-related financial risks and the central bank's policy response: Literature review [J]. Qinghai Finance*, 2022 (06): 26-33.
- [2] Jiang Jingqiang. *The challenge and response to the Financial system [J]. Hainan Finance*, 2022 (04): 29-35.
- [3] Liu Bolin, Li Linlin. *International practice and experience of transforming finance from the perspective of climate change risk [J]. Financial Development Review*, 2022 (03): 54-69.
- [4] Gao Rui, Wang Ying, Cao Tingqiu. *Evidence on Climate Change and Macro Financial Risk — from 58 representative countries worldwide [J]. Nankai Economic Studies*, 2022 (03): 3-20.
- [5] Liu Guiping. *Efforts were made to improve the climate risk management capacity of the financial system [J]. China Finance*, 2022 (05): 9-11.
- [6] Zhang Shuai, Lu Liping, Zhang Xingmin, Wang Hui. *Assessment, pricing, and policy response of climate risks in the financial system: A literature-based review [J]. Financial Review*, 2022, 14 (01): 99-120 + 124.
- [7] Chang Yanqiu. *Analysis of the correlation between climate change risk and financial risk [J]. Northern Finance*, 2021 (12): 47-49.
- [8] Wang Xin. *Green Finance Development and Climate Risk Management [J]. Financial Economy*, 2021 (07): 3-9.