Sentiment Analysis on Climate Change for Sustainable Investment

Keane Ong
Asian Institute of Digital Finance
National University of Singapore
keane.ongweiyang@u.nus.edu

Rui Mao
College of Computing and Data Science
Nanyang Technological University
rui.mao@ntu.edu.sg

Deeksha Varshney
College of Design & Engineering
National University of Singapore
deeksha2@nus.edu.sg

Ranjan Satapathy
Institute of High Performance Computing
Agency for Science, Technology and Research
satapathy_ranjan@ihpc.a-star.edu.sg

Erik Cambria
College of Computing and Data Science
Nanyang Technological University
cambria@ntu.edu.sg

Gianmarco Mengaldo*
College of Design & Engineering
National University of Singapore
mpegim@nus.edu.sg

***Abstract*—This research involves examining whether sentiment on climate change can be accounted as a systematic risk factor within sustainable finance. Tweets related to climate change from 2014-2022 are collected via the Twitter API. RoBERTa is then deployed to label the collected tweets according to sentiment polarity. Thereafter, DistilBERT is utilised to classify tweets according to their different climate change topics *Aggregated, Impact, Politics & Policy, Mitigation, Rootcause*. Using these sentiment and topic labels, monthly sentiment scores are aggregated from the frequency of sentiment polarity (number of positive, and negative tweets), and segmented according to the topics that the tweets are classified under. On the other hand, we collect the monthly returns of various green ETFs, alongside ETFs that contain stocks from other sectors (technology, oil and gas), as well as 2 baskets of stocks (sustainable and heavy-polluting). We run Fama-French regressions by conducting time series regression of monthly returns on the traditional 3 factors, with specific addition of sentiment scores.**

I. Introduction

Given the pressing issues of climate change, sustainable finance, or investment that considers sustainability factors, has become an important topic [1], [2]. In fact, sustainable investing can be a significant driver of emissions reductions and climate change mitigation [3]. On the other hand, machine learning and statistics have been used on data to inform investment decisions [4]. For example, promise has been shown in the use of natural language techniques to compute public sentiment for trading and equity investment [5]–[8]. They can provide valuable insight into market sentiment or specific equities and stocks [9]–[11].

Considering these developments, this paper aims to use machine learning and regression techniques to investigate the relationship between climate change sentiment and returns in green or sustainable assets (ETFs, Stocks). The study focuses on tweet content and funds originating in the United States due to its large number of Twitter users, frequent natural disasters, and the presence of listed sustainable ETFs with significant value. Thereafter, from deriving sentiment scores and segmenting them into their various topics, we incorporate them into a modified Fama-French model to investigate whether sentiment scores can account for an additional source of market risk associated with green asset returns.

*corresponding author

II. Related Work

To account for the returns on sustainable assets, different studies within empirical asset research have opted to modify existing factor models to incorporate climate risk. For instance, [12] opted to derive a carbon risk premium by taking the difference of average returns between pollutant and green stocks worldwide, while [13] constructed a green-minus-brown factor by taking the difference between high climate performance companies and low climate performance companies. While these studies have constructed climate change risk factors in the same spirit as Fama and French [14] (by taking return differences between different assets), they do not significantly control for multi-collinearity. Specifically, although green stocks may have a premium over brown stocks because of differences in climate performance, there may also be interdependence between climate performance and the size, book-market ratio, investments, and profitability of firms.

To reduce multi-collinearity to a greater extent within a risk-factor regression, alternative methods (apart from taking the return difference for traditional (SMB, HML, CMA, RMW) factors) should be explored. For instance, the climate risk factor can be constructed directly from textual content, facilitated by natural language processing. In fact, textual analysis has already shown promise for extracting important information from climate change-related text corpora [15]. [16] implemented a textual analysis of newspapers coupled with a mimicking portfolio approach to construct climate change hedge portfolios.

Moreover, in finance, NLP methods such as twitter sentiment analysis provide useful information for stocks [17]. For example, researchers have computed polarity scores by labelling financial news according to sentiment, with these scores used to predict stock prices via deep learning [18]. Studies have also included integrating artificial neural networks with a sentiment polarity index to predict investment

2375-9259/24/$31.00 ©2024 IEEE
DOI 10.1109/ICDMW65004.2024.00067

trends [19]. These approaches have been applied to sustainable finance. Particularly, [20] employed lexicon-based sentiment analysis to predict stocks prices for companies with sustainability interests.

However, a limited number of studies are relevant to understanding the dynamics between climate change sentiment and sustainable investing beyond stock forecasting. [21] used StockTwits to derive climate change sentiment and thereafter its relationship with an Emission-Minus-Clean stock portfolio. To the best of my knowledge, there has not been a study that attempts to utilise natural language processing methods to derive price climate change sentiment as a systematic risk factor for returns on green assets. To further explore this area, this paper focuses on the sentiment around climate change on Twitter and its relation to clean energy investment.

III. Data

A. Twitter API

For the collection of tweets related to climate change, the official Twitter API is utilised via academic research access. The keyword of 'climate change' is employed for dates from 2014 to 2022, with the tweet location set as the US. The time period in question is decided due to the increased interest in sustainable finance from the mid 2010s onward.

B. Labelled Tweets Dataset

A separate labelled dataset from [22] is utilised to train, evaluate, and test both the sentiment and topic classifier. This dataset contains 2312 climate change related tweets labelled with the sentiment polarity *positive, negative, neutral* as well as the different topics *impact, mitigation, politics and policy, others*. The labelled dataset is split into training, validation, and test data in the proportion of 0.64, 0.16 and 0.2 respectively. This labelled dataset has been expertly annotated by human evaluators, and will serve as the ground truth. It will be utilised to evaluate the accuracy of the sentiment and topic classification algorithm.

C. Collection of Asset Data

Financial data for monthly asset returns are collected for ETFs corresponding to different sectors (Clean energy, oil & gas, market-wide) from 2014 to 2022. Clean energy ETFs include: *Share Global Clean Energy ETF (ICLN), Invesco Wilderhill Clean Energy ETF (PBW)* Oil & gas ETFs include: *iShares US Oil & Gas Exploration & Production ETF (IEO), Energy Select Sector SPDR Fund (XLE)* Market-wide technology ETFs include: *Invesco QQQ Trust Series 1 (QQQ).*

The motivation for studying ETFs is due to their ability to replicate benchmark indexes that track a diverse range of companies situated in a particular sector. In other words, ETFs can be a useful representation of companies in a specific industry that is related or unrelated to sustainable finance. Despite these advantages ETFs, the underlying characteristics of ETFs may make it a demanding task for their return patterns to be captured by systematic factors. For example, their prices can deviate from their net asset values, perhaps due to liquidity effects, which may not be captured by traditional Fama-French factors. To elaborate, ETF's unique mechanism involving redemption of units helps to keep ETF prices close to NAV, whereby differences in ETF price and NAV can be arbitraged away through authorised participants. When ETF prices are lower than NAV, these participants can buy ETFs, redeem them for their individual stocks, and sell the individual stocks to make the difference, causing a downward pressure on the stock prices such that ETF NAV more closely match ETF prices. However, sustained deviations can arise due to illiquidity, which makes it difficult to buy and sell assets to arbitrage away the difference between ETF NAV and prices. Alongside this is also how ETF structures can be complexly structured to generate returns, which may reduce their coherence with market factors.

As such, to control for these underlying weaknesses, we also include the asset returns of 2 different basket of stocks related to sustainable and heavy-polluting sectors respectively, with each basket corresponding to 50 stocks each. The green (sustainable) basket includes the following industries and corresponding stocks (represented by their tickers): Renewable Energy & Equipment: TSLA, FSLR, ALB, NEE, BEP, RUN, CWEN, CSIQ, SPWR, ORA, TERP, AY, AZRE, JKS, PEGI, DQ, BE, HASI, REGI, MDDNF. Sustainable Agriculture & Food: BYND, CVGW, VRYYF, AMRS, SFM, APPH, FDP, UNFI, LWAY, HAIN, DAR, INGR, BG, ADM, TATYY. Green Transportation: NIO, PLUG, WKHS, BLNK, CHPT, QS, RIDE, PTRA, HYLN, FSR, GP, XL, LEV, SOLO, GOEV. Clean Technology & Equipment: ITRI, ENPH, TRMB, APTV, BLDP, FCEL, OLED, ECL, IEX, DHR.

On the other hand, the brown (heavy-polluting) basket includes the following: Oil & Gas Exploration and Production: XOM, CVX, RDS.A, BP, COP, OXY, EOG, MRO, APA, PXD, DVN, HAL, SLB, TOT, XEC. Coal Mining: BTU, ARCH, ARLP, CEIX, CTRA, HCC, NRP, SXC, METC, HCC. Chemicals: DOW, DD, EMN, LYB, EMN, FMC, PPG, APD, SHW, CE. Steel & Iron: MT, X, NUE, STLD, CLF, PKX, CMC, GGB, RS, VALE. Utilities (Fossil Fuel-Based): DUK, SO, D, AEP, PPL.

D. Data Processing

Tweets are processed for sentiment classification by replacing weblinks and Twitter usernames by 'http' and '@user', respectively. This omits unnecessary information such as user and website names, enabling the classifier to learn weblink and username features more consistently.

IV. Sentiment Analysis

A. RoBERTa Model

A pre-trained RoBERTa model from Huggingface, Cardiffnlp (twitter-roberta-base-sentiment) [23] is used for sentiment analysis on the climate-change related tweets. The RoBERTa has advantages over other BERT models. It is more accurate and robust as it is trained on larger datasets. Additionally, RoBERTa's training procedure is also an improvement to that of the standard BERT. Specifically, the 'next sentence prediction' job is eliminated, and dynamic masking is used to alter the masked tokens over training epochs. Finally, the Cardiffnlp RoBERTa is pre-trained on fifty-eight million tweets and is finetuned for sentiment analysis using the TweetEval benchmark. It can detect informal language used in tweets such as emoticons.

B. Model Training

For training, the sequences that are fed into the RoBERTa (text from tweets) are transformed into embedding vectors, with each vector being mapped to a single word in the sequence. The transformer encoder uses a self-attention mechanism to create contextual embeddings and learns the context of each word. To represent the semantic information in the tweet, the contextual embeddings for each word are joined into a single vector. The pre-trained RoBERTa on the sentiment detection task is then fine-tuned by adding a classification layer at the end of the feature extractor model to forecast the appropriate sentiment class *positive, negative, and neutral* corresponding to each input sequence (tweet text).

C. Transfer Learning

Further transfer learning is done to optimise the RoBERTa model for sentiment classification of climate change related tweets. In essence, transfer learning has been proven to improve the accuracy of BERT algorithms for domain-specific language and tasks, by further fine-tuning the neural network weights to the task in question [24].

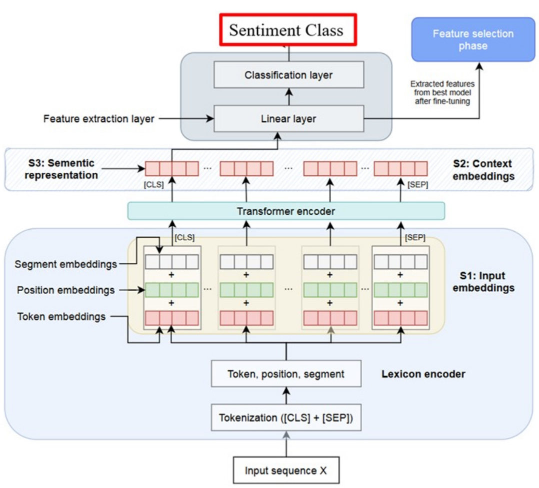

Fig. 1: RoBERTa Transformer Architecture

D. DistilBERT Model

It involves extending the training of an already pre-trained algorithm on the new task at hand. For our application, we extend the training of the algorithm to sentiment classification on a climate tweets dataset in section III-B. The training parameters include training on 3 epochs, a training batch size of 16, evaluation batch size of 64 and a weight decay of 0.01. Transfer learning improves the test set classification accuracy from 70% to 78%.

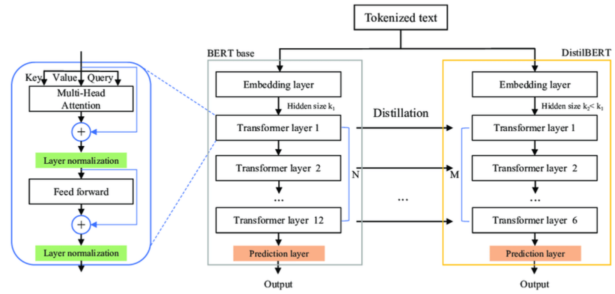

Fig. 2: DistilBERT Transformer Architecture

V. Topic Classification

Topic classification methods are also employed to classify the collected tweets into different topics: *Impact, Mitigation, Politics and Policy, Others, Rootcause*. Essentially, we employ DistilBERT [25] to classify the tweets because it promises to be lighter and quicker (40% less parameters) compared to typical BERT models while retaining a significant accuracy rate (97% performance).

A. Model Training

DistilBERT is trained in similar fashion as RoBERTa in section IV-B. However, the classification layer in our DistilBERT use-case forecasts the topic class *impact, mitigation, politics and policy, others, rootcause* instead of sentiment polarity, from the text input sequence. Apart from this, DistilBERT also does away with token-type embeddings, pooling capabilities, and fewer layers are present within its architecture relative to RoBERTa. Our DistilBERT attains a roughly 70% accuracy for topic classification on the previously mentioned labelled climate change dataset III-B.

VI. Further Appreciation of Classification Accuracy

As mentioned, the RoBERTa model is able to accurately classify the sentiment of tweets (positive, negative, neutral) approximately 8 out of 10 times (78%), while DistilBERT is able to accurately classify the topic of tweets (*Politics & Policy, Mitigation, Impact, Rootcause, Others*) roughly 7 out of 10 times (70%). *Politics & Policy* refer to tweets related to the discussion of government, politics,and policies surrounding climate change, *Mitigation* refers to tweets discussing current efforts and potential ways to mitigate climate change, *Impact* discusses the consequences of climate change, while *Rootcause* discusses the different causes of climate change. *Others* constitute climate tweets which do not fall under the mentioned categories. To provide a further appreciation of tweets that are classified positive and negative, alongside their different topics, as shown by table I.

A. Sentiment Scores

While we can measure both positive and negative sentiment, this study will focus on negative sentiment, given how literature has shown its relatively more substantial impact on asset

TABLE I: Examples of tweets and their associated classifications from BERT

Tweet	Sentiment	Topic
When it comes to climate change, expensive gas is a good thing. It makes you think twice about driving anywhere and adding to the problem with your emissions. It also makes you think about trading in your gas guzzler for an electric.	Positive	Mitigation
Elon Musk is trying to be responsible by creating a solution to hinder climate change by putting his ideas into his Teslas which also means saving fossil fuels for times when electric energy isn't the answer...does not take a genius to know that	Negative	Mitigation
We are here with a panel of experts to discuss how coastal communities in North Carolina can become more resilient to future storms and climate change	Positive	Impact
Our democracy is hanging by a thread, women's rights are imperiled, climate change threatens us, global authoritarianism is raging.	Negative	Impact
Today I set out how UK is helping small island developing states future proof against climate change and extreme weather by investing in resilient infrastructure. PM has announced £11.6bn over the next 5 years to help tackle climate change around the world.	Positive	Politics & Policy
True, but it would solve their global warming/climate change agenda. Really put a damper on the green new deal, and end the fossil fuel debate.	Negative	Politics & Policy
Not Gung Ho overall the common misperceptions about climate change, e.g. the world literally ending within a handful of years, but I heavily support the advent of electric vehicles, because, climate change aside, ICE cars are massively polluting. #GoTesla #DamnTheAutoDealers	Positive	Rootcause
Hey Joe you could have Fauci concoct another virus and a more lethal one and kill off alot more Americans and that would bring down our carbon footprint on the ozone. Climate change is not the most pressing issue to the world.	Negative	Rootcause
In other news, the weather is quite delightful today. Looking forward to frolicking around NYC this afternoon because we know climate change won't let us be great all week.	Positive	Others
Nobody really is afraid of climate change except very old people who watch the news all day and people of Twitter.	Negative	Others

prices as compared to positive sentiment [26]. Additionally, we only focus on tweets that are classified from selected topics that are most meaningful (Politics & Policy, Mitigation, Impact, Rootcause). On top of these topics, we have also aggregate all the tweets collected without discrimination of their separate topics under *Aggregated*. We compute monthly climate change sentiment, $\boldsymbol{x}$, corresponding to each of the different topics *Aggregated, Politics & Policy, Mitigation, Impact, Rootcause* from minimum and maximum normalisation of the monthly frequency of negative labelled tweets from 2014 to 2022. To qualify this with an example, climate change sentiment scores for the topic *Mitigation*, is derived from collecting tweets for each month that are classified as relevant to the discussion on climate change mitigation, and thereafter obtaining the monthly frequency of negatively labelled tweets corresponding to this subset of tweets. Thereafter, we apply normalisation to these frequencies. The same procedure is repeated for each of the topics. The procedure for our normalisation is straightforward and explained by

$$\boldsymbol{x} = 10 * \frac{x - x_{min}}{x_{max} - x_{min}}, \tag{1}$$

where x_{max} & x_{min} are the maximum and minimum monthly frequency of negative labelled tweets across 2014 to 2022. We multiply the result of our minimum and maximum normalisation by a factor of 10, to give us climate change sentiment scores, $\boldsymbol{x}$, in the range of 0 to 10. The motivation for doing so is to ensure that the scale of $\boldsymbol{x}$ coincides with the scale of other factors in our regression, such as SMB, HML, $R_m - R_f$. This importantly enhances the interpretability of the coefficient of $\boldsymbol{x}$ upon conducting the regression which will be later explained.

VII. Fama-French 3-Factor Regression

The Fama-French three-factor model was introduced by Eugene Fama and Kenneth French in the early 1990s [14]. The model was developed to explain anomalies (deviations from expected model outcomes) in the CAPM, which could not account for certain market behaviors. Specifically, the Fama-French model aimed to explain why stocks with low market capitalizations (small firms) and stocks with high book-to-market ratios (value stocks) tended to outperform the market. The model is explained by

$$R_a - R_f = \alpha + \beta_1 \left(R_M - R_f\right) + \beta_2 SMB + \beta_3 HML_t + \varepsilon \tag{2}$$

Market Factor $R_M - R_f$ is the excess return of the market over the risk-free rate. It represents the systematic risk associated with the market. Size Factor (SMB) or "Small Minus Big" represents the difference in returns between small-cap and large-cap stocks. A positive SMB value indicates that small-cap stocks have outperformed large-cap stocks. Value Factor (HML): stands for "High Minus Low". It represents the difference in returns between stocks with high book-to-market ratios (value stocks) and those with low book-to-market ratios (growth stocks). A positive HML value indicates that value stocks have outperformed growth stocks.

The Fama-French model suggests that, in addition to the market risk, the size and value factors also play a significant role in determining stock returns. The model has been widely adopted in empirical finance research and has been extended in various ways, including the addition of investment and profitability factors in later versions such as [27]. Other extensions included the addition of a momentum factor [28]. While these models have showed significant performance and accuracy in accounting for market-wide risk factors as a source of excess returns, we investigate whether factors more specific to the asset in question (climate change sentiment for green assets) can function as additional risk factors for returns. Specifically, we run the following regression:

$$R_a - R_f = \alpha + \beta_1 \left(R_M - R_f\right) + \beta_2 SMB + \beta_3 HML_t + \beta_4 \boldsymbol{x} + \varepsilon \tag{3}$$

The extended Fama-French model includes an additional factor, $\boldsymbol{x}$, which is the negative climate change sentiment scores corresponding to the different topics as explained earlier. We evaluate whether climate change sentiment accounts for additional risk. The motivation for extending the 3-factor Fama-French model instead of the 5-factor Fama-French model, is to ensure that the new added climate change sentiment factor $\boldsymbol{x}$ functions as a source of risk that is relatively more isolated from the other factors within the model. To elaborate, we hypothesise that within the context of a 5-factor model, the additional risk factors of CMA (conservative minus aggressive investment) or RMW (robust minus weak profitability) while conceptually distinct from climate change sentiment, may have potential interactions with climate change sentiment. To explain further, companies that invest aggressively in new technologies might be doing so to address climate change, perhaps borne out of the influence of climate change sentiment (specific views toward climate change). This may intertwine CMA and climate change sentiment. Similarly, green companies may have profitability resulting from sales driven by the climate change sentiment of customers (i.e. investments in solar energy), linking RMW and climate change sentiment. Consequently, climate change sentiment may not constitute a unique contribution to returns within the context of of an extended five factor model.

To qualify, the same rationale may also be applied to climate change sentiment and the different factors of size and book market ratio. That is, the interactions between climate change sentiment and size, book market ratio, may mean that within an extended 3-factor model, climate change sentiment may also not constitute an isolated unique contribution to returns.

However, while we cannot eliminate multi-colinearity entirely, we can reduce it by extending the 3 factor model in place of the 5 factor model, thereby reducing the number of factors that may interact with climate change sentiment.

VIII. Results & Discussion

A. Monthly Sentiment Scores Over Time

To understand how climate change negative sentiment scores vary over time, we plot the monthly scores from 2014 to 2022. These monthly scores correspond to the tweets belonging to different topics, as mentioned previously, such as *Aggregated, Mitigation, Politics and Policy, Root Cause, Impact*. The plot is shown in Figure 3

Evidently, the monthly scores follow different trends for different topics, with each score peaking at different time periods. These different trends will yield different results in our modified Fama-French regressions following the Eq. (3).

B. Constants and Coefficients of Regression

We run the regressions according to Eq. (2) and (3). In our tables, we will highlight the results of our regressions, with particular attention to the values of α, β_4, and the respective t values as well as R^2. To qualify, we are running the regression (2) without specific addition of the climate change sentiment factor $\boldsymbol{x}$ to compare with (3). Additionally, the first column depicts the asset that we are running the regression for. We construct multiple tables to represent our results, each referring to the different topic of tweets *aggregated, mitigation, politics & policy, root cause, impact* corresponding to our $\boldsymbol{x}$. For example, the table for *mitigation* refers to the regression (3) run with negative climate change sentiment scores derived for tweets which have been topically classified as describing climate change mitigation, whereas *aggregated* refers to climate change sentiment scores derived for all tweets. For the assets, GB and BB stands for Green Basket and Brown Basket respectively, while the ETFs are denoted by their tickers.

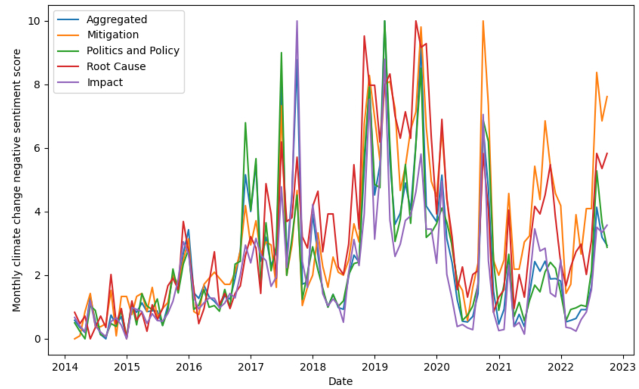

Fig. 3: Monthly climate change negative sentiment across 2014 to 2022

To evaluate whether $\boldsymbol{x}$ can function as a systematic risk factor, we should ideally observe that the magnitude of α should decrease with the inclusion of $\boldsymbol{x}$ into the Fama-French 3 factor regression, particularly for sustainable/ green ETFs and stocks. This would mean that the risk factor would account for some of the excess returns, and thus possibly explain some of the asset returns. However, besides observing α, we also look at the coefficients of regression, particularly β_4, as well as the error and t value statistics to examine statistical significance of the new regression. As we run the traditional Fama-French 3 factor regression on the different assets, as shown in table II, we realise that the calculated t-values for α are mostly below the acceptable threshold of 2.0, except for the Green Basket. A smaller t-value can have several interpretations, specifically, that there is less evidence that the dependent variable is non-zero when all the predictive variables are zero.

However, we are cautious in interpreting lower t-values for α as a measure of the other independent variables being better able to explain returns. Instead, this analysis will focus more on the t-values of the coefficients, alongside the R^2 values, as well as some insight into the magnitude of α

Firstly, the calculated R^2 values for the different regressions (2) & (3), with and without the inclusion of the sentiment scores remains fairly constant throughout, indicating that the

TABLE II: FF3 factor regression

Asset	α	β_1	β_2	β_3	R^2	$t(\alpha)$	$t(\beta_1)$	$t(\beta_2)$	$t(\beta_3)$
GB*	0.96	1.25	1.21	-0.54	0.10	2.46	14.5	8.17	-5.14
BB**	0.34	1.21	0.37	0.85	0.26	1.78	27.7	5.08	16.1
ICLN*	0.11	1.04	0.34	-0.33	0.45	0.20	7.81	1.54	-2.06
PBW*	-0.11	1.34	1.16	-0.55	0.59	-0.16	8.86	4.57	-2.95
IEO**	-0.14	1.41	0.74	1.34	0.69	-0.22	9.78	3.07	7.67
XLE**	-0.26	1.18	0.43	1.17	0.71	-0.53	10.7	2.34	8.73
QQQ†	0.23	1.12	-0.24	-0.32	0.94	1.71	36.6	-4.68	-8.52

* Sustainable ETF/Stocks
** Heavy-polluting ETF/Stocks
† Technology-dominated ETF (Nasdaq)

TABLE III: FF3+1 regression for *Aggregated* tweets

Asset	α	β_1	β_2	β_3	β_4	R^2	$t(\alpha)$	$t(\beta_1)$	$t(\beta_2)$	$t(\beta_3)$	$t(\beta_4)$
GB	0.88	1.25	1.21	-0.53	0.03	0.10	1.47	14.5	8.16	-5.11	0.17
BB	1.35	1.20	0.39	0.84	-0.40	0.26	4.60	27.6	5.26	15.9	-4.54
ICLN	-0.44	1.05	0.33	-0.33	0.22	0.46	-0.50	7.82	1.50	-2.01	0.84
PBW	-1.51	1.36	1.14	-0.53	0.57	0.60	-1.52	9.04	4.53	-2.90	1.88
IEO	1.20	1.40	0.76	1.33	-0.54	0.70	1.28	9.81	3.19	7.67	-1.90
XLE	0.59	1.17	0.44	1.16	-0.34	0.72	0.81	10.67	2.42	8.70	-1.55
QQQ	0.56	1.12	-0.23	-0.32	-0.13	0.94	2.82	37.2	-4.67	-8.78	-2.23

TABLE IV: FF3+1 regression for *Politics & Policy* tweets

Asset	α	β_1	β_2	β_3	β_4	R^2	$t(\alpha)$	$t(\beta_1)$	$t(\beta_2)$	$t(\beta_3)$	$t(\beta_4)$
GB	0.63	1.26	1.21	-0.53	0.13	0.10	1.04	14.5	8.06	-5.06	0.70
BB	1.24	1.20	0.40	0.84	-0.36	0.27	4.18	27.4	5.46	15.9	-3.97
ICLN	-0.67	1.05	0.31	-0.32	0.32	0.46	-0.76	7.88	1.41	-1.99	1.16
PBW	-1.75	1.37	1.10	-0.5215	0.67	0.61	-1.75	9.14	4.39	-2.86	2.16
IEO	1.22	1.39	0.78	1.32	-0.55	0.70	1.28	9.75	3.29	7.62	-1.88
XLE	0.61	1.17	0.46	1.16	-0.35	0.72	0.83	10.6	2.50	8.66	-1.54
QQQ	0.54	1.12	-0.23	-0.32	-0.12	0.94	2.64	36.7	-4.51	-8.76	-2.00

TABLE V: FF3+1 regression for *Mitigation* tweets

Asset	α	β_1	β_2	β_3	β_4	R^2	$t(\alpha)$	$t(\beta_1)$	$t(\beta_2)$	$t(\beta_3)$	$t(\beta_4)$
GB	0.70	1.26	1.21	-0.53	0.07	0.10	1.02	14.4	8.17	-5.13	0.46
BB	1.03	1.20	0.38	0.85	-0.20	0.26	3.07	27.3	5.11	16.14	-2.50
ICLN	-1.00	1.06	0.34	-0.33	0.33	0.46	-1.02	7.95	1.53	-2.06	1.40
PBW	-2.07	1.39	1.15	-0.54	0.59	0.61	-1.86	9.20	4.61	-2.99	2.17
IEO	0.41	1.40	0.74	1.34	-0.16	0.69	0.38	9.63	3.07	7.64	-0.63
XLE	0.23	1.17	0.43	1.17	-0.15	0.71	0.28	10.5	2.34	8.70	-0.73
QQQ	0.49	1.12	-0.24	-0.32	-0.08	0.94	2.17	36.4	-4.68	-8.58	-1.43

addition of climate change sentiment scores has minimal impact on regression model fit.

Out of the green assets (GB, ICLN, PBW), we observe that only for GB, does the regression (3) with the inclusion of the added sentiment risk factor yield lower α relative to the original FF3 factor regression (2). Moreover, these results are only limited to negative climate change sentiment, $\boldsymbol{x}$, derived from the topics *Aggregated, Politics & Policy, Mitigation*.

To elaborate, from tables III, IV & V which each correspond to regression (3) ran on tweets from the topics *Aggregated, Politics & Policy, Mitigation*, where α values are at 0.88, 0.63 & 0.70 respectively. This is lower compared to the α value derived from the FF3 factor regression 2 which is 0.96 as seen from table II. This may indicate that the addition of climate change sentiment $\boldsymbol{x}$ has accounted for some of the excess returns of GB.

Given the short time frame of the study, we can only hypothesise possible risk-based explanations describing the interaction of climate change negative sentiment, $\boldsymbol{x}$, and asset returns for GB, based on the results that we have obtained. Specifically, sustainable investments, at least for green basket stocks GB, have positive risk exposure to negative climate sentiment scores, $\boldsymbol{x}$ derived from tweets corresponding to the topics *Aggregated, Politics & Policy, Mitigation*.

Moreover, we can also attempt to hypothesise why these results are specific to these topics instead of *Impact, Root cause*. Tweets from the topics of *Politics & Policy, Mitigation*, involve discussions on the political & non-political solutions for climate change. As such, green assets yield greater returns with rising negative climate change sentiment from tweets discussing climate solutions, perhaps because they are seen as riskier as they are more exposed to climate sentiment

TABLE VI: FF3+1 regression for *Impact* tweets

Asset	α	β_1	β_2	β_3	β_4	R^2	$t(\alpha)$	$t(\beta_1)$	$t(\beta_2)$	$t(\beta_3)$	$t(\beta_4)$
GB	1.11	1.25	1.22	-0.54	-0.06	0.10	1.87	14.5	8.18	-5.14	-0.31
BB	1.20	1.21	0.38	0.84	-0.40	0.27	4.15	27.7	5.17	15.8	-3.97
ICLN	-0.24	1.04	0.34	-0.33	0.17	0.46	-0.28	7.79	1.52	-2.01	0.56
PBW	-1.15	1.35	1.15	-0.53	0.50	0.60	-1.18	8.92	4.55	-2.86	1.44
IEO	0.99	1.41	0.75	1.32	-0.54	0.69	1.07	9.84	3.14	7.59	-1.64
XLE	0.41	1.18	0.44	1.16	-0.32	0.71	0.58	10.69	2.38	8.63	-1.26
QQQ	0.63	1.12	-0.24	-0.33	-0.19	0.94	3.26	37.81	-4.76	-9.00	-2.80

TABLE VII: FF3+1 regression for *Root cause* tweets

Asset	α	β_1	β_2	β_3	β_4	R^2	$t(\alpha)$	$t(\beta_1)$	$t(\beta_2)$	$t(\beta_3)$	$t(\beta_4)$
GB	1.00	1.25	1.21	-0.54	-0.01	0.10	1.48	14.38	8.15	-5.11	-0.08
BB	1.20	1.19	0.36	0.84	-0.25	0.27	3.65	27.20	4.86	15.74	-3.21
ICLN	-0.58	1.05	0.35	-0.32	0.21	0.46	-0.59	7.85	1.59	-1.96	0.88
PBW	-1.74	1.38	1.18	-0.51	0.50	0.60	-1.58	9.11	4.71	-2.78	1.84
IEO	0.05	1.41	0.73	1.40	-0.06	0.69	0.05	9.64	3.03	7.57	0.22
XLE	-0.12	1.18	0.43	1.17	-0.04	0.71	-0.14	10.53	2.31	8.61	-0.21
QQQ	0.49	1.12	-0.24	-0.32	-0.08	0.94	2.20	36.43	-4.79	-8.68	-1.47

relating to these topics. Increased negativity on these topics may constitute growth in scepticism of climate solutions provided for by green assets, elevating the risk of holding them. Thus, higher returns are yielded in compensation. However, to qualify, this explanation is not vindicative, it is instead an attempt at explaining how climate change sentiment may drive green asset returns. Just like the profitability factor in the Fama-French 5 factor model, the actual risk-based explanation may remain obscure and requires more research.

Moreover, our provided analysis must include a critique of t values associated with the regressions. Specifically, although the t-values of the other coefficients β_1, β_2, β_3 remain fairly constant and statistically significant between the FF3 factor regression (2) and the aforementioned regressions (3), the t-value of β_4 for the regressions run on *Aggregated, Politics & Policy, Mitgation* remains below the threshold of 2 as shown from tables III, IV & V. Thus, the coefficient β_4 may not be statistically significant in accounting for the excess returns of GB. On the other hand, we can analyse the coefficients of β_4 to examine the economic significance of the factor, in place of its statistical significance. Even though we have adjusted the scale of the sentiment scores from (1), we observe that the coefficients of β_4 for GB are minuscule for the topics *Aggregated, Mitigation*, < 0.10, as shown from tables III & V, while the coefficient is relatively more significant for *Politics & Policy* at 0.13 from table IV.

To add further context to our analysis, we have opted to conduct the regressions for a mixture of different assets, including stocks and ETFs which are heavy polluting (BB, IEO, XLE) and are more unrelated to climate change (QQQ which is primarily technology-dominated). We observe that the α increases for regressions which include the negative climate change sentiment, following Eq. (3), from tables IV, III, VI, V, as compared to the original FF3 regression (2) as shown in table II. This indicates that the modified model (3) is unable to account for the excess returns of these non-green assets at least when including climate sentiment factor $\boldsymbol{x}$ derived from the topics *Politics & Policy, Aggregated, Impact, Mitigation*, more so than the original FF3 model (2).

The exception is the regressions (3) of XLE corresponding to tweets from the *Mitigation* topic, as well as the regressions (3) of XLE & IEO corresponding to tweets from the *Root cause* topic. While the α value of regression (2) for XLE & IEO, as seen from table II is -0.26 & -0.14 respectively, α is -0.12 & 0.05 for XLE & IEO corresponding to the regression (3) for tweets derived from the topic *Rootcause*, and α is 0.23 for XLE for regression (3) for tweets derived from *Impact*. Importantly, although a lower α may indicate (subject to statistical significance), a greater ability of the modified model to explain excess returns, a change in sign of α indicates a new interpretation of α. A change from negative α to positive α, indicates an over performance relative to the benchmark risk factors in place of an under performance. This raises an inquiry on the validity of the new benchmark, which arises from the specific addition of negative climate change sentiment to the traditional 3 factors, relative to the old benchmark (3 factors), at least for the discussed non-green assets (XLE, IEO). Moreover, the results of these regressions of (3) corresponding to XLE & IEO for *Rootcause, Impact* are still statistically insignificant, as they yield <2 for the t value of β_4. It is also difficult to interpret the negative β_4 values for the regression (3) for *Rootcause, Impact*, which is indicative of a negative risk exposure to negative sentiment regarding these topics, going against intuition.

Interestingly, the values of the coefficients β_4 corresponding to the regressions of the various non-green assets are negative for climate change sentiment corresponding to each of the different topics, while the β_4 coefficients remain positive for green assets. Although the values of α do not reduce for regressions involving ICLN & PBW, and only reduce for the aforementioned regressions involving GB, the positive coefficients of β_4 imply positive risk exposure to negative climate sentiment for the different green assets studied.

IX. Conclusion

All in all, in comparison with the original Fama-French model (2), the regressions of the modified Fama-French model with the inclusion of climate change sentiment (3) show more promise for accounting for the excess returns for green assets (specifically GB) relative to other non-green assets. For instance, the derived α is of lower magnitude with the inclusion of the additional climate change sentiment factor. However, the derived results are not statistically significant, given that the t values for this regression fall below the threshold of 2. This is despite our best attempts at mimicking the factor of climate change sentiment, with the implementation of sentiment analysis to classify sentiment alongside topic classification to classify the discussions.

The limitations of the study are realised. Specifically, although text content from Twitter has the potential to account for asset returns [29], Twitter sentiment related to climate change and its various topics *Aggregated, Politics & Policy, Mitigation, Impact, Root Cause*, have limited influence on asset returns, at least in the context of a risk-factor within a modified Fama-French factor regression. For future work, a different approach to factor construction for sentiment might yield better results. Specifically, instead of mining text data related to the climate change topic generically, it may be more relevant to collect textual data more specifically related to the assets. This can be done by adding queries which combine specific financial keywords and the topic of climate change. This way, the tweets collected will be more associated to the financial aspects of climate change, making them more relevant to asset returns. Additionally, alternative text mining techniques such as metaphorical analysis [30] and neurosymbolic AI [15], can also provide a fresh perspective on the textual data related to the assets. Explainable AI techniques can also be utilised to derive meaningful financial insights [31], [32], with examples such as [33], [34] highlighting methods for more robust and granular data analysis.

References

[1] T. Stobierski, "What is sustainable investing?," 2022.

[2] K. Ong, R. Mao, R. Satapathy, E. Cambria, J. Sulaeman, G. Mengaldo, *et al.*, "Explainable natural language processing for corporate sustainability analysis," *arXiv preprint arXiv:2407.17487*, 2024.

[3] R. Bakken, "What is sustainable investing and why is it important," 2021.

[4] S. Morris, "A quick guide to data-driven investing," 2022.

[5] Y. Ma, R. Mao, Q. Lin, P. Wu, and E. Cambria, "Quantitative stock portfolio optimization by multi-task learning risk and return," *Information Fusion*, vol. 104, p. 102165, 2024.

[6] I. Chaturvedi, Y.-S. Ong, I. Tsang, R. Welsch, and E. Cambria, "Learning word dependencies in text by means of a deep recurrent belief network," *Knowledge-Based Systems*, vol. 108, pp. 144–154, 2016.

[7] E. Cambria, B. Schuller, B. Liu, H. Wang, and C. Havasi, "Statistical approaches to concept-level sentiment analysis," *IEEE Intelligent Systems*, vol. 28, no. 3, pp. 6–9, 2013.

[8] K. Du, F. Xing, R. Mao, and E. Cambria, "Financial sentiment analysis: Techniques and applications," *ACM Computing Surveys*, vol. 56, no. 9, pp. 1–42, 2024.

[9] K. Ong, W. van der Heever, R. Satapathy, E. Cambria, and G. Mengaldo, "Finxabsa: Explainable finance through aspect-based sentiment analysis," in *2023 IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 773–782, IEEE, 2023.

[10] E. Cambria, X. Zhang, R. Mao, M. Chen, and K. Kwok, "SenticNet 8: Fusing emotion AI and commonsense AI for interpretable, trustworthy, and explainable affective computing," in *International Conference on Human-Computer Interaction (HCII)*, 2024.

[11] E. Cambria, *Understanding Natural Language Understanding*. Springer, ISBN 978-3-031-73973-6, 2024.

[12] U. S. Akbar, N. A. Bhutto, and S. K. O. Rajput, "Impact of climate change hazard: Carbon risk on stock returns and asset pricing," 2021.

[13] M. Rostad and M. A. Myking, "Climate change: the transition risk: an empirical analysis of the inclusion of a green-minus-brown factor in common factor models," Master's thesis, 2020.

[14] E. F. Fama and K. R. French, "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, vol. 33, no. 1, pp. 3–56, 1993.

[15] C. Duong, V. C. Raghuram, A. Lee, R. Mao, G. Mengaldo, and E. Cambria, "Neurosymbolic ai for mining public opinions about wildfires," *Cognitive Computation*, vol. 16, no. 4, pp. 1531–1553, 2024.

[16] R. F. Engle, S. Giglio, B. Kelly, H. Lee, and J. Stroebel, "Hedging Climate Change News," *The Review of Financial Studies*, vol. 33, pp. 1184–1216, 02 2020.

[17] H. K. Sul, A. R. Dennis, and L. Yuan, "Trading on twitter: Using social media sentiment to predict stock returns," *Decision Sciences*, vol. 48, no. 3, pp. 454–488, 2017.

[18] Y. Peng and H. Jiang, "Leverage financial news to predict stock price movements using word embeddings and deep neural networks," *arXiv preprint arXiv:1506.07220*, 2015.

[19] S. K. Khatri and A. Srivastava, "Using sentimental analysis in prediction of stock market investment," in *2016 5th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*, pp. 566–569, IEEE, 2016.

[20] D. V. Gakhar and S. Kundlia, "Impact of sentiments on stock returns, volatility and liquidity," *International Journal of Economic Policy in Emerging Economies*, vol. 14, no. 5-6, pp. 536–565, 2021.

[21] C. Santi, "Investor climate sentiment and financial markets," *International Review of Financial Analysis*, p. 102490, 2023.

[22] C. Duong, Q. Liu, R. Mao, and E. Cambria, "Saving earth one tweet at a time through the lens of artificial intelligence," in *2022 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–9, 2022.

[23] F. Barbieri, J. Camacho-Collados, L. Neves, and L. Espinosa-Anke, "Tweeteval: Unified benchmark and comparative evaluation for tweet classification," *arXiv preprint arXiv:2010.12421*, 2020.

[24] M. Mozafari, R. Farahbakhsh, and N. Crespi, "A BERT-based transfer learning approach for hate speech detection in online social media," in *Complex Networks 2019: 8th International Conference on Complex Networks and their Applications*, pp. 928–940, 2019.

[25] V. Sanh, L. Debut, J. Chaumond, and T. Wolf, "Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter," *arXiv preprint arXiv:1910.01108*, 2019.

[26] S. Kelly and K. Ahmad, "Estimating the impact of domain-specific news sentiment on financial assets," *Knowledge-Based Systems*, vol. 150, pp. 116–126, 2018.

[27] E. F. Fama and K. R. French, "A five-factor asset pricing model," *Journal of Financial Economics*, vol. 116, no. 1, pp. 1–22, 2015.

[28] E. F. Fama and K. R. French, "Size, value, and momentum in international stock returns," *Journal of financial economics*, vol. 105, no. 3, pp. 457–472, 2012.

[29] E. Bartov, L. Faurel, and P. S. Mohanram, "Can twitter help predict firm-level earnings and stock returns?," *The Accounting Review*, vol. 93, no. 3, pp. 25–57, 2018.

[30] R. Mao, Q. Lin, Q. Liu, G. Mengaldo, and E. Cambria, "Understanding public perception towards weather disasters through the lens of metaphor," in *Proceedings of the thirty-third international joint conference on artificial intelligence*, 2024.

[31] G. Mengaldo, "Explain the black box for the sake of science: Revisiting the scientific method in the era of generative artificial intelligence," *arXiv preprint arXiv:2406.10557*, 2024.

[32] W. J. Yeo, W. van der Heever, R. Mao, E. Cambria, R. Satapathy, and G. Mengaldo, "A comprehensive review on financial explainable ai," *arXiv preprint arXiv:2309.11960*, 2023.

[33] J. Wei, H. Turbé, and G. Mengaldo, "Revisiting the robustness of post-hoc interpretability methods," *arXiv preprint arXiv:2407.19683*, 2024.

[34] H. Turbé, M. Bjelogrlic, C. Lovis, and G. Mengaldo, "Evaluation of post-hoc interpretability methods in time-series classification," *Nature Machine Intelligence*, vol. 5, no. 3, pp. 250–260, 2023.