Nostradamus: Weathering Worth

Sunday, 17th October, 2021

Alapan Chaudhuri
IIIT Hyderabad
alapan.chauhduri@research.iiit.ac.in

Zeeshan Ahmed
IIIT Hyderabad)
zeeshan.ahmed@research.iiit.ac.in

Shivansh Subramanian
IIIT Hyderabad
shivansh.s@research.iiit.ac.in

Shreyas Pradhan
IIIT Hyderabad
shreyas.pradhan@research.iiit.ac.in

***Abstract*—Nostradamus, inspired by the French astrologer and reputed seer, is a detailed study exploring relations between environmental factors and changes in the stock market. This paper describes our analysis system using historical stock market data, historical climate data, and various climate indicators to represent the correlation and causation between them accurately. We have conducted our study based on the US market, climate, and environment, which reliably shows a significant relationship between climate and stock market changes. Lastly, we have taken four natural disasters as a case study to observe the effect they have on people's emotional state and their participation in the stock market.**

I. Introduction

A stock market is a aggregation of buyers and sellers of stocks. A popular saying which goes with the stock market is that the people decide the worth of the company. Though it aptly describes the definition, it is often seen that such is not the case, where due to various reasons companies, expected or not, end up deviating from market trends. Our goal here is to establish a relation between climate conditions, weather and disasters with respect to such deviations, so that we can predict and analyse such events in the future and take appropriate actions.

II. Dataset

A. Data Extractions

We have taken historical stock data (all the way till 1980s) for multiple companies using Yahoo Finance API and ZEPL US Stock Market Data. We have also extracted sustainability data about each company from Yahoo for general usage and insights. All of the data is publicly available for free, and is also stored on our GitHub. For Environmental Data we have used KNOEMA Environmental Data Atlas to get environmental data for multiple countries over many years. We have got 40+ parameters related to CO2 emissions, Fuel consumption, Use of nitrogen amongst others per country per year. To get climate related data we have used NOAA Climate Data, which gives us temperature, precipitation and snowfall of a given area per day.

B. Data Cleaning and Enhancements

Once we have obtained the data requires as data-frames, we performed feature selection to reduce the noise in the data. This made our process more efficient and decreased space limitations. We have performed two different analysis: yearly, and daily. For yearly data analysis, we have grouped the weather data, climate data and historical OHLC data by year to aptly represent the columns. For daily data analysis, we also computed Moving Averages, Volume Weighted Average Price, Uncertainty and 50 Day Standard Deviation as well.

$$VWAP = \frac{\sum_0^{n-1} Volume_i \times \frac{High+Low+Close}{3}}{\sum_0^{n-1} Volume} \tag{1}$$

$$SD = \sqrt{\frac{\sum_0^{n-1} \mid x - \bar{x} \mid}{n}} \tag{2}$$

We have stored all the data required in separate 'csv' files so as to decrease the runtime. We also created a pipeline which enables the user to easily fetch all the data for any stock they require.

III. Environmental Factors

Our hypothesis was that there are multiple environmental factors which effect the government and the people directly, but not the companies so much. Companies change their policies, which changes their revenues, based on reaction from the government and their customers. In a graph, this would be reflected by an initial change in trend (based on people's opinion), which later settles a bit and then changes based on whether or not government changes their policies.

$$r_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \sum \bar{y})^2}} \tag{3}$$

A. Method

We define the value of correlation r_{xy} as the ratio of sum of deviations divided by root of product of sum of square of deviations. A huge value of correlation index (close to 1) indicates that the environmental factor and the company's

profits share a causal relationship, whereas a correlation index closer to -1 represents an inverse of such relationship.

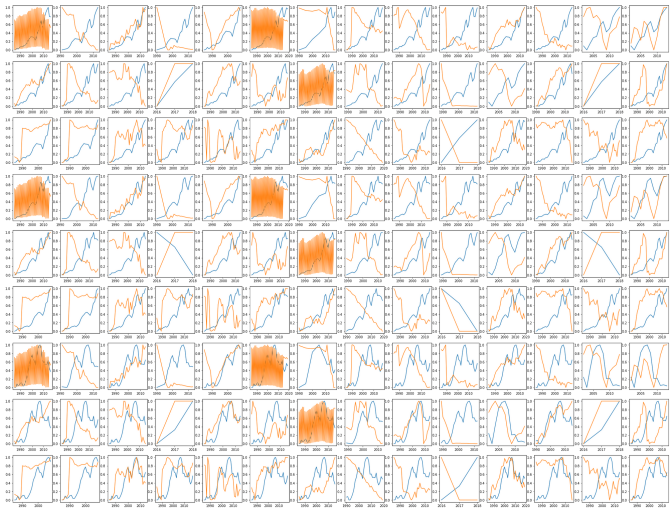

Fig. 1. Represents correlation between various companies' stock prices vs co2 emission in USA over the years. A bigger plot is available on our github

B. Inferences

Multiple stocks show different correlation between co2 emissions and stock prices, and each of them signify important information related to their sector.

1) Strong Correlations: High positive value of correlation usually mean that there is an interdependence (Strong Correlation) between a stock and an environmental factor. Consider the stock (British Petroleum Company), which is an oil and gas company. It is obvious that as the company's production has a direct effect on CO2 emission. Hence, we can confidently infer that CO2 emission values are effected by BP's stock.

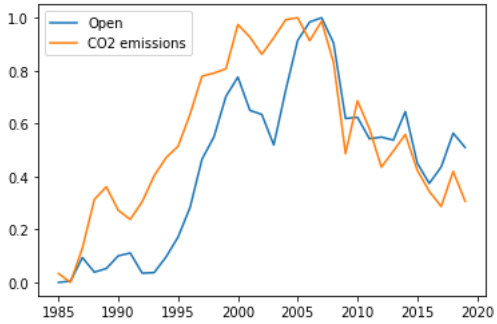

Fig. 2. Shows the correlation between open prices of BP over the years vs CO2 emissions in USA. A correlation value as high as 0.8159 was observed

2) Causation versus Correlation: It is not always true that a high value of correlation means that there is an interdependence between the stock prices and CO2 emissions. Consider the stock of AAPL (Apple) and the CO2 emissions from gaseous fuel consumption. The correlation value observed was close to 0.93. We know that Apple (a tech company), is obviously not dependent on CO2 emissions from gaseous fuel consumption. Inspite of that, it has a high correlation with that factor. This is not a result of a dependence between the two things. The high correlation is simply a coincidence as CO2 emmissions from gaseous fuels are rising because of the rapid population growth and because nuclear and other clean energy sources are not very prevalant. Hence Correlation is not always a result of Causation.

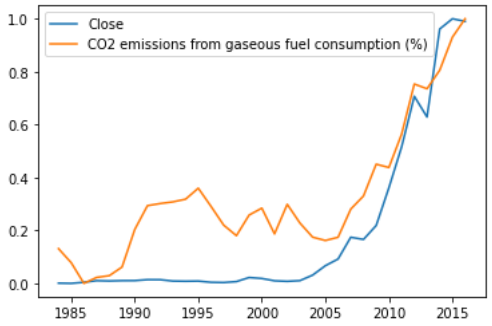

Fig. 3. Shows the correlation between open prices of AAPL over the years vs CO2 from gaseous fuel consumption in USA. A correlation value as high as 0.93 was observed

3) Hidden Correlations: Sometimes a stock and an environmental factor seemingly unrelated end up having a high correlation. We observed that there are 2 possible explanations for it: either it is just a coincidence (as shown above), or it has a hidden correlation. As an example, consider the stock of EOD (Wells Fargo Global) vs CO2 emissions in USA. The graph has a high positive correlation value of 0.927. However, it is likely that there are hidden correlations due to dependencies of carbon emissions on an industry which also determines whether the prices of companies which own/invest in the same. This includes large companies such as large banks and firms. Hence, this high correlation is very likely not a result of coincidence but a result of purposeful investing/decisions taken by the company.

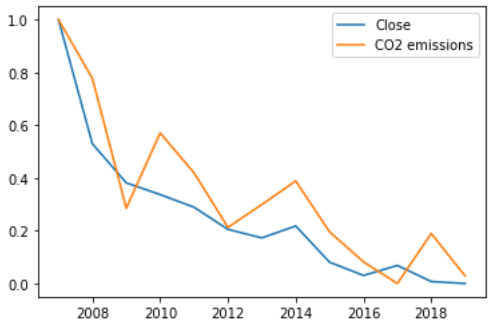

Fig. 4. Shows the correlation between open prices of EOD over the years vs CO2 emissions in USA. A correlation value as high as 0.927 was observed

4) Low Positive Correlations: A low positive correlation usually means that the stock and the environmental factor in consideration is independent of each other. Consider the case of XOM (Exxon Mobil) and Agricultural Methane Emissions, which has a low correlation of 0.234. Hence, we can most

of the times conclude that these two are independent of each other. In rare cases though, they could be dependent and still have a low correlation.

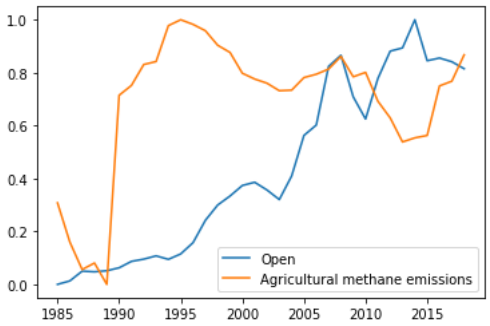

Fig. 5. Shows the correlation between open prices of XOM over the years vs Agricultural Methan Emissions in USA. A correlation low correlation value of 0.234 was observed

5) High Negative Correlations: Highly negative values of correlation generally imply an inverse effect between the company's production/success and the factor in consideration. Consider the volume of the stock CVX (Chevron Corporation) and the factor Terrestrial and marine protected areas. This has a highly negative correlation of -0.899. It is fair to assume that when the number of terrestrial and marine protected areas increases, the volume of the stock CVX, which is an energy industry decreases. Hence, a highly negative correlation value means inverse dependence.

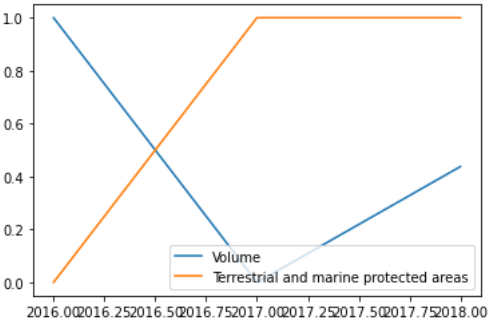

Fig. 6. Shows the correlation between open prices of CVX over the years vs Terrestrial and Marine Protected area in USA. A correlation high negative correlation value of -0.899 was observed

C. Predictions

The global emission levels are predicted to be stranded due to growing awareness about climate change and a noticeable switch to renewable sources as a primary energy sources. As we earlier saw the high correlation between the carbon emission levels and the stock price of BP, it can be analyzed that in coming times the stock price of BP and other large oil companies will begin to fall. This also means that there will be a rise in the stock prices of companies that provide an alternative source of fuels such as solar panels and wind mills.

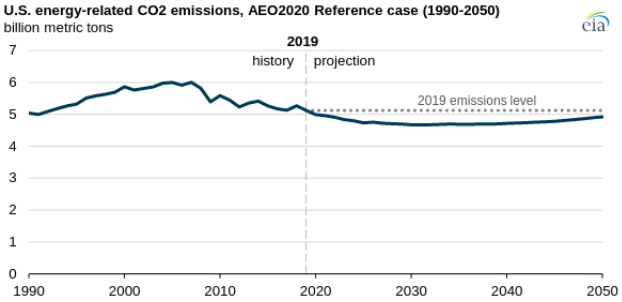

Fig. 7. US Energy-related CO2 emissions

D. Industrial Revolution

The example of correlation and not causation can further be understood by considering the rise in the population and workforce. The rise in population meant more consumption of energy and since we lack any large source of energy other than fossil fuels it means that there will be a rise in the pollution levels.

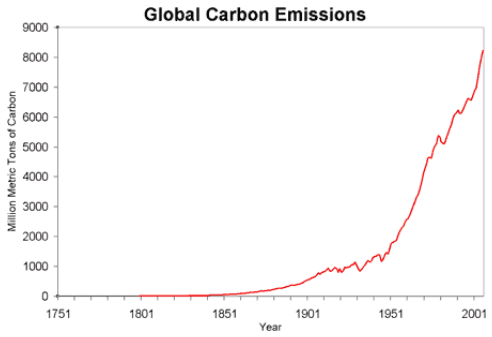

Fig. 8. Global carbon emissions

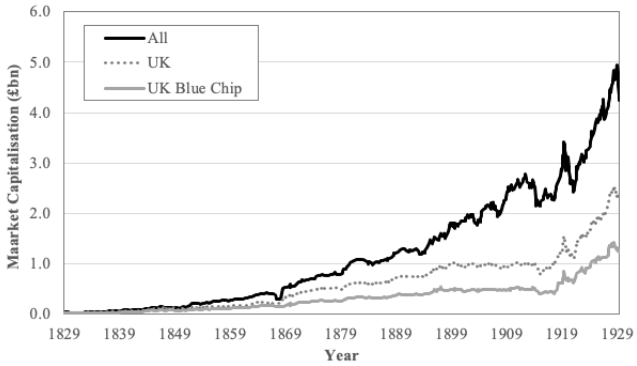

Fig. 9. Market Capital Growth

If the stock history of the industrial revolution times along with the carbon emission levels of the same was available, it can be seen that the rising workforce and demand directly meant a rise in carbon levels. This is a correlation due to common sources of the rise.

IV. Natural Disasters

As an experiment, we performed case studies on four natural or man-made disasters occurring in USA and how the stock prices reacted to it.

A. California Wildfires

In terms of property damage, 2017 was the most destructive wildfire season on record in California at the time, surpassed by only the 2018 season, with a total of 9,560 fires burning. Throughout 2017, the fires destroyed or damaged more than 10,000 structures in the state (destroyed 9,470, damaged 810).

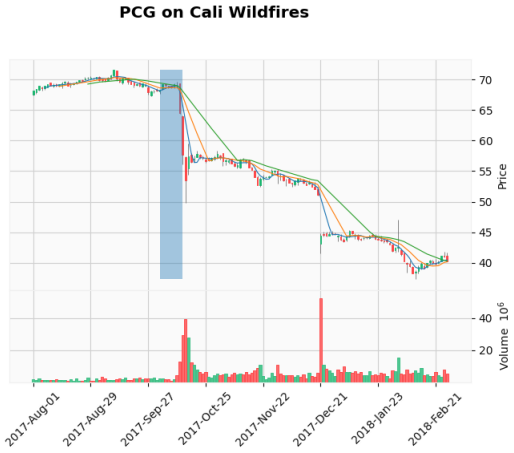

Fig. 10. Stock prices of PCG during the 2017=18 California wildfires. Tl blue bar represents when the wildfires started

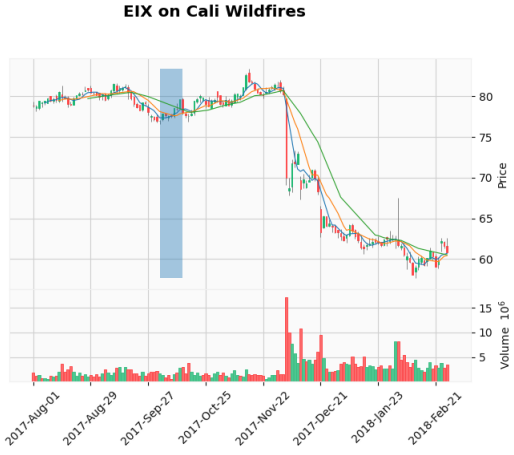

Fig. 11. Stock prices of EIX during the 2017=18 California wildfires. The blue bar represents when the wildfires started

We can observe an interdependence between the graphs in this case. The stock of PCG which is a gas and electric company was affected right after the disaster took place. But now consider EIX. Its stock fell sharply some time after PCG's. We know that EIX depends on PCG for its production. This is also evident through the graph as its stock was not affected directly after the wildfires but a few days later. This gives clear information about dependence of companies on each other.

B. Texas Storm

In February 2021, the state of Texas suffered a major power crisis, which came about as a result of three severe winter storms sweeping across the United States. The storms caused a massive electricity generation failure in the entire state leading to shortages of water, food, and heat. More than 4.5 million homes and businesses were left without power, at least 210 people were killed directly or indirectly with some estimates going as high as 702. The Texas Grid failure was majorly caused by the inadequately winterized natural gas equipment. NRG suffered the most as can be seen in the following plots. It can also be emphasized that the analysis has to be done on a small time scale as companies often bounce back from such losses over a longer period of time such as a year.

Fig. 12. Stock prices of NRG during the 2021 texas winter storm. The blue bar represents when the winter storm was at its peak

C. Katrina Hurricane

In August 2005, Katrina was a category 5 Atlantic hurricane that caused over 1,800 deaths and $125 billion in damage. This damage was mainly focused on New Orleans and the surrounding areas. Large transportation companies such as C.H. Robinson had to face a heavy loss due to denial of services. Since hurricanes leave a relatively smaller effect on companies, they are able to resume their services. so the loss is made back.

As is seen here, companies related to travel and transport profited here due to the urgent requirement of vehicles for relief and rescue.

D. 9/11

Although 9/11 is not a natural disaster, it is worth considering how it affected the stock market. The effect is very visible on the entire stock market due to direct or indirect loss. As seen in the bottom graphs, the prices go down relatively fast after the event took place. Due to its surprising nature, the impact's nature and magnitude were no less than a natural disaster over the country.

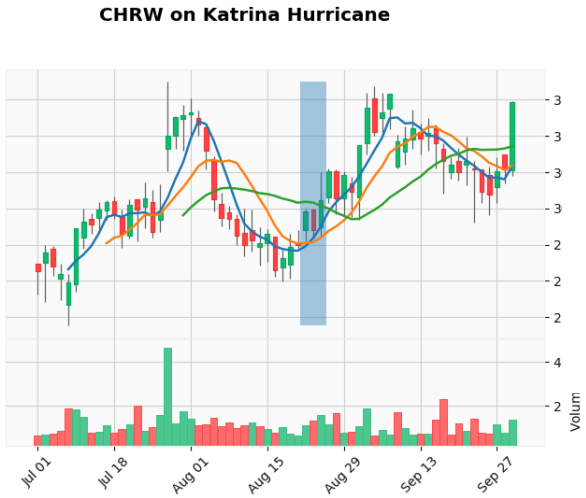

Fig. 13. Stock prices of CHRW during the 2005 Katrina Hurricane. The blu bar represents when the hurricane occurred

Fig. 14. Stock prices of KMX during the 2005 Katrina Hurricane. The blue bar represents when the hurricane occurred

Fig. 15. Stock prices of ALK during 9/11. The blue bar represents when the event occurred

Fig. 16. Stock prices of DIA during 9/11. The blue bar represents when the event occurred

Fig. 17. Stock prices of SPY during 9/11. The blue bar represents when the event occurred

For all these disasters, stocks of companies that seemed related to those events were taken into consideration. We observer that AMK, a company related to airlines, suffered huge losses. Even stock indices suffered gap losses representing the general mood of market and traders after the disasters when everything seemed unsure.

In agreement to our hypothesis, the stocks of those companies suffered a heavy hit after those events.

V. Weather

We try to observe correlation between daily weather and daily trades. These two have a correlation of 0.0338, which can be analyzed as no correlation It can be seen that there is no correlation between volume and precipitation. Due to the availability of the stock market electronically, trades can be made from anywhere. This pattern is new, but before the availability of the stock market on electronics, the trades had to be done in a building to which people traveled.

VI. Future Work

The data has been analyzed so far which has us given some insights, it can be used to build upon for prediction and more analysis for determining the effect of stocks on future natural events.

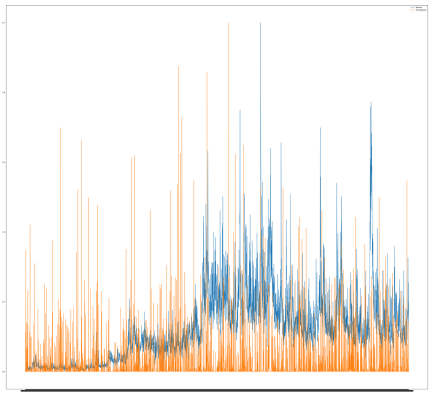

Fig. 18. No particular correlation between weather and daily trades. Orange represents daily precipitation and blue represents volume of stocks traded

Once the prediction model has been built, it can be used to analyze environmental factors. For example, when we compare the predicted value of a stock to its actual value, and there is a big difference at some point, it implies that some kind of environmental factor affected that stock at that point.

Events such as hurricanes and heatwaves recur periodically. Due to rising climate change, they will only become more frequent. This means that past analysis on such events will help us determine the kind of impact it will happen on the companies it had an impact on.

VII. Conclusion

This analysis has revealed that there are correlations underlying between the stock value and the environmental factors such as weather and natural disasters. Further, major events which were unpredictable but aren't natural causes also have a heavy effect on the market. It can also be seen that the increase in supply and demand had a direct impact on the market stock, along with which there was an impact on environmental factors such as carbon emissions.

There are cases that have been explained where the impact of a natural disaster first hit the company which had a direct link to the resources that were lost followed by a loss in the companies that depended upon the primary companies after a delay.

The data can be understood and analyzed with graphs and other tools, predictions can be made with the help of machine learning models which could be trained on used in a case such events are repeated again.

The data has been analyzed so far which has us given some insights, it can be used to build upon for prediction and more analysis for determining the effect of stocks on future natural events.

Once the prediction model has been built, it can be used to analyze environmental factors. Example, when we compare the predicted value of a stock to its actual value, and there is a big difference at some point, it implies that some kind of environmental factor affected that stock at that point.

Events such as hurricanes and heatwaves recur periodically. Due to rising climate change, they will only become more frequent. This means that past analysis on such events will help us determine the kind of impact it will happen on the companies it had an impact on.

References

1) Environmental Disasters and Stock Market Performance by Ishuwar Seetharam (Stanford University, 2017)
2) The impact of disasters and terrorism on the stock market by Tchai Tavor and Sharon Teitler-Regev
3) Do Earthquakes Shake Stock Markets? by Susana Ferreira and Berna Karali
4) Researchers Find Weather May Have An Effect On The Stock Market (Forbes Advisor Report)
5) The Effect of Weather on Stock Trading by Dasol Jeong (University of California, Berkeley)