Handelshøyskolen BI

GRA 19703 Master Thesis

Thesis Master of Science 100% - W

Predefinert informasjon

Startdato:	09-01-2023 09:00 CET	**Termin:**	202310
Sluttdato:	03-07-2023 12:00 CEST	**Vurderingsform:**	Norsk 6-trinns skala (A-F)
Eksamensform:	T		
Flowkode:	202310||11184||IN00||W||T		
Intern sensor:	(Anonymisert)		

Deltaker

Navn:	Nicholas Algie

Informasjon fra deltaker

Tittel *:	How to price weather derivatives and model temperature? What are the effects of global warming on temperature forecasting?
Navn på veileder *:	Yifan Zhu

Inneholder besvarelsen konfidensielt materiale?:	Nei	**Kan besvarelsen offentliggjøres?:**	Ja

Gruppe

Gruppenavn:	(Anonymisert)
Gruppenummer:	329
Andre medlemmer i gruppen:	Deltakeren har innlevert i en enkeltmannsgruppe

How to price weather derivatives and model temperature? What are the effects of global warming on temperature forecasting?

AUTHOR

Thesis Supervisor
Yifan Zhu

BI Norwegian Business School

Hand-in date:
03.07.2023

Campus:
BI Oslo

Examination code and name:
GRA 19703 Thesis Master of Science

Program:
Master of Science in Business – Major in Sustainable Finance

Acknowledgments

The master thesis was written as a part of the Master of Science double degree program in between BI Norwegian Business School and Politecnico of Milan in business with a major in sustainable finance.

The thesis was written during the academic year 2022/2023 semester by Nicholas Algie. The work is independent and original.

We thank our thesis advisor, Yifhu Zan, Associate Professor –
Department of Finance at BI Norwegian Business School, for his continuous support, advice, and constructive feedback throughout the overall process. His guidance has been helpful. We would also like to thank our friends and families for their support.

ABSTRACT

This thesis investigates the complex relationship between the pricing of temperature weather derivatives with its underlying temperature, with a focus on how temperature can be modeled and the consequences of global warming in this niche options market.
The research studied the different models available on the subject with the goal of finding a pricing model that fits temperature, and all the risks attached to it, into the pricing model. The classic option pricing models, such as Black and Scholes can't be adopted as a key condition of their success is the negotiable nature of the underlying.

Looking at what the literature offers, there are valid mathematical pricing models. However, they are very complicated, and not every organization that wants to adopt these hedging tools has the knowledge and capital to comprehend them. This is the reason the paper tries to use the payoff system, suggested on the Chicago Mercantile Exchange, the main source of the options with some modifications to make the pricing fit for programming. The paper adopts time series methods of research and is all carried out on R.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0 INTRODUCTION

The research question that the paper will try to answer is:

How to price weather derivatives and how the temperature of these can be modeled to be fit for forecasting? What are the effects of global warming on these?

The thesis will use Data from Options in the United States and the weather temperature data is precise and available for most of the zones of the States; therefore, the research will be carried out through actual options prices, simulated prices through the model and take into consideration global warming and the recent events extreme on how the temperature is modeled.

We expect that global warming does affect the modeling and, therefore, ohe pricing and the way these financial derivatives are adopted. After a first theoretic section, where the paper will go through all the necessary notions and information that are required to comprehend the research, the thesis includes a thorough literature review analyzing the existent literature, research papers, and articles that can contribute to the success and quality of this paper.

A central piece of the paper consists of the data retrieving and sources, and before starting the methodology section of the research there will be a description of the reasons way the data used was sourced and how it was sourced.

Once the methodology of the research is outlined, the results will be analyzed, and the interpretations carried out. The conclusion will be anticipated by some final considerations on the future of such derivatives, using the results found, and how this paper contributes to the few researchers that are out there, including how it can be implemented in a broader entrepreneurial context.

2.0 WEATHER DERIVATIVES

Before initiating the research topic explanation and the literature review that will start the process of researching whether global warming has affected the pricing model of weather derivatives or/and has affected how this financial asset is adopted.

Weather derivatives are recent, launched in 2001 (Wolfram Schlenker, Charles A. Taylor 2020), and have been a relatively growing market, hedging financial assets that some companies have at their disposal to hedge the risks related to their business operations and the weather. If all the companies in the world were able to perfectly hedge these risks, there would be a total savings of 250 million US dollars (Don Cyr, Martin Kusy 2010). The peculiarity of these options is that the underlying contract is not based on a price of a specific material or an exchange rate but is based on specific pre-determined data; therefore, the use of these derivatives must be for a risk that does not have direct monetary value (Andreas Mueller and Marcel Grandi, 2002). The most popular data on which weather derivatives are based are temperature and the amount of rain (liters) in a season. In this paper, the focus will be more on temperature derivatives. It is important to state that these options are designed for hedging the risk, therefore any type of speculative position that a financial takes with taking on these options is not meant it changes the purpose of why these commodities were made.

The commodities with temperature data as the underlying are the most popular in the US (Andreas Mueller and Marcel Grandi 2002), which is where the research will be carried out. The options are designed for specific sectors that carry great risk related to factors that companies can't control. The most famous example is the one related to the ice cream industry, if the temperature is high the consumer won't be as price sensitive and will consume ice cream even at a greater price, the ice cream can only forecast how a season could be (in terms of degrees) but if the forecasts are wrong and the pricing strategy is a failure the ice-cream company will incur in a loss that is not dependant to their strategy unless they hedged. The industries and the sector that purchases the most weather commodities are the energy sector and the agricultural one.

The pricing model how is now composed of a simple structure based on a base temperature value of 65 degrees Fahrenheit (18 degrees Celsius), this temperature is the standpoint from which temperature fluctuations are considered Heating Degree Days (HDD) and Cooling Degree Days (CDD).

HDD and CDD are indices that show the average temperature for a given period, using 65 degrees standpoint. The degree days are taken by doing the difference, in absolute values, between the basic temperature and the actual temperature.

CDD and HDD calculating method:

$$CDD = \sum_{t=1}^{n} (T(t) - 65) \quad \forall T(t) > 65 \ \ (1)$$

$$HDD = \sum_{t=1}^{n} (65 - T(t)) \quad \forall T(t) < 65 \ \ (2)$$

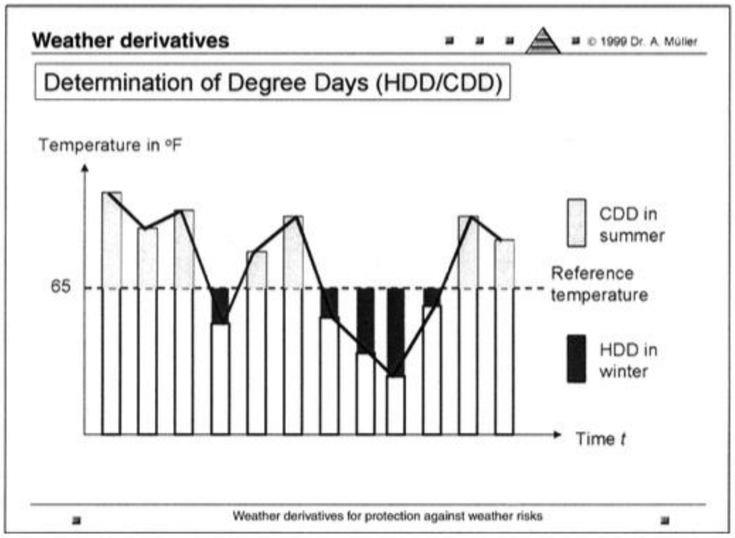


Figure 1: Heating Degree Days (HDD) and Cooling Degree Days (CDD)

These two measures are for the two halves of the year, where HDD concerns days during the winter season and CHD concerns days during the summer season. These derivatives work exactly as classic options, as there is a strike price where you can apply call and put options, plus all the strategy combinations available for options, e.g., straddles, and collars; are applicable for these weather options. HDD options give hedging opportunities against risks related to worm winters and CDD against cold summers.

Table 1:
System for temperature options

Option type	Protection against ...	Exercised when	Payout
HDD call	... overly cold winters	HDD > strike value	tick size × (HDD − strike value)
HDD put	... overly warm winters	HDD < strike value	tick size × (strike value − HDD)
CDD call	... overly hot summers	CDD > strike value	tick size × (CDD − strike value)
CDD put	... overly cool summers	CDD < strike value	tick size × (strike value − CDD)

Table 1: System for temperature options

As is shown in table 1, there is a predetermined strike price k, which in this case is a value since the underlying of the options is temperature, over or under the HDD and CDD must go to activate the hedging payout. This payout is defined by the tick size y, a predefined amount of dollars that corresponds with a single degree of temperature above or under the strike value (table 1). Under a formula format the payoff of a call HDD for example would be as follows (Ali, Z., Hussain, J. & Bano, Z. 2023):

$Payoff = y \max\{H - k, 0\}$ (3)

This pricing method is called the Burning Cost premium, which this the only applicable one as for these options the underlying is temperature and is not negotiable, a key characteristic for the more familiar option pricing method Black-Scholes. The main issue, the one on which some of the focus of our research will be is the fact that the Burn Cost premium does not take into consideration the fluctuation trends of the weather and is not protected by the consequences of global warming and the consequences to all the GHG emissions (Andreas Mueller and Marcel Grandi, 2000). It has been shown (Campbell, Dieboldt (2005)) that when dealing with daily temperature the errors of the daily estimate accumulate (as the cumulative sum would suggest) and can impact greatly the forecast and consequentially the payouts, as the period of the data increases the error also increases. Furthermore, when we look at the formula for calculating the HDD or CDD, key measures for these options, it is noticeable that there is a sum of each day's temperature above the standard 65. This is the main reason why the pricing is not protected by the unpredictability of the weather, which has increased drastically in the last twenty years, as there could be a short period of incredibly hot winter during a cold season. This event may compromise the value of the option and cause a loss in function of the financial hedging, as energy may receive the revenue from their main operations (average cold winter) plus the payout that is caused by the

warm period, as the sum is not weighted based on the days but just one the total temperature.
There is also a risk concerning the future, we know are aware of the increase in temperature that the globe is facing: +1.5 to 2 degrees. It is relevant to ask the question of how this would affect these options and if the changes made until now are valid enough to face these challenges.

To further comprehend, we take in consideration the case example presented from the Chicago Mercantile Exchange on their website.
The example considers a timeframe from May to September, which represents a 5-month "summer" seasonal strip. The seasonal 10-year average for New York is stated as 1352. The utility's objective is to protect against potential losses due to mild weather. The strike price is set at 1250, with a cap at 1050. The option premium is $600 per contract, and the maximum potential loss (cost of "The Hedge") is $1,020,000. The maximum potential payout is $6,800,000.

The strategy calculation involves determining the number of contracts needed for effective hedging. Based on historical averages, the calculation is as follows: $5,780,000 divided by 170 (10 ticks * $20/tick) equals 34K per CDD. Dividing 34K per CDD by $20 gives 1,700 contracts. Multiplying 30 ticks, $20 per tick, and 1,700 contracts results in a cost of $1,020,000. Therefore, the utility would pay a premium of $1,020,000 for $5,780,000 in protection. (Weather Options Overview – CME Group.)

3. MODEL COMPREHENSION: USEFUL NOTIONS

In this section, the paper will discuss and illustrate useful notions that the reader must comprehend before going into specific sources.

Global Warming

Since the industrial revolution in 185o the GHG emission in terms of flow emitted, starting by the European western country to the industrialization of China, this increase in production has caused the start of global warming. Global warming can be defined as the increase in temperature, over a period of thirty years or more, in the surface and sea surface area due to human activity. However, is key to understanding how emissions work to consider that 58% of the stock of CO2 in the air was accumulated between 1950 and 1958 (IPCC report 2022) and is the stock that has impacted the most in the last thirty years, instead of the current flow of emission will impact the future impact the next thirty years and that's why is important to reduce the flow of emissions.

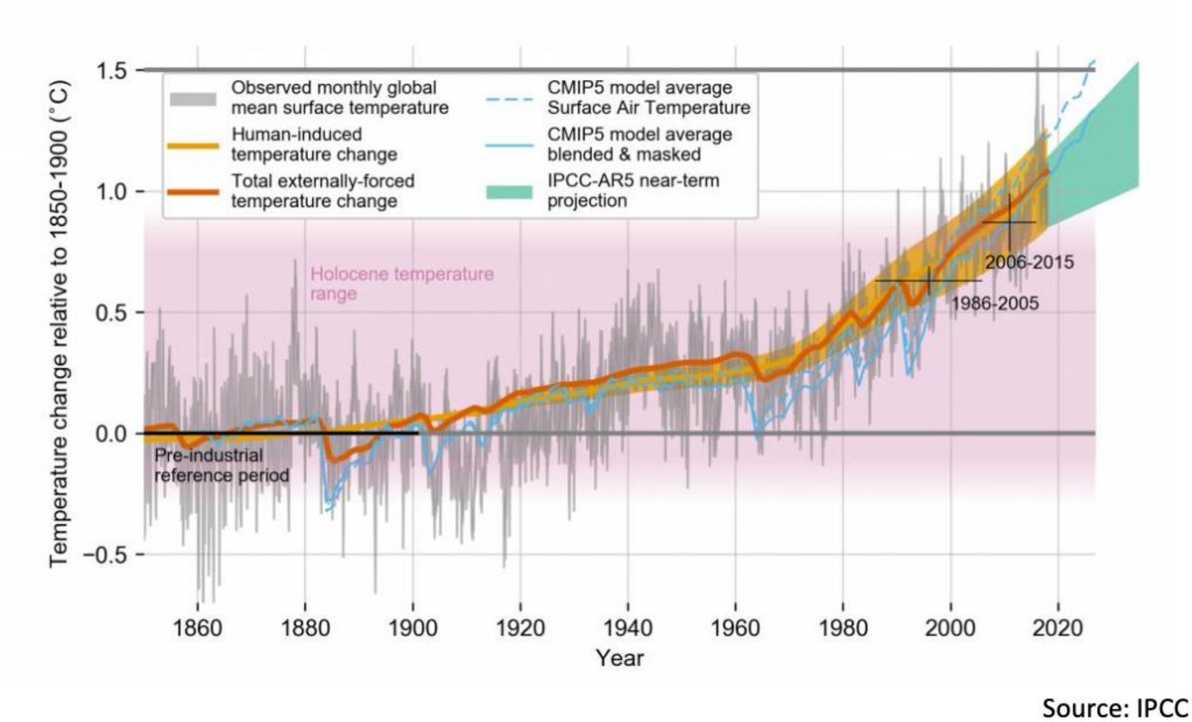


Figure 2: Temperature change relative to years 1860 to 2020

As it is possible to see from the table and as it is clearly stated in the last IPCC report there has been a fearful increase in temperature since the pre-industrial levels, approximately 1°C, and the risk is to reach 2°C or above is concerning, which will have a severe consequence on the planet as we know it. The table also shows a projection that is now tackled by all the restrictions and objectives that the EU is trying to implement and have been discussed at the recent COP conference in Egypt.

Global warming has also been an important recent topic of discussion in the finance sector, the issue is whether environmental events have a real effect on market fluctuations. There is no proper answer if the environmental and social concerns have value from a strictly financial perspective. However, this research can show empirically if this temperature increases and the increase in temperature fluctuations in a single season directly affect a type of financial investment available to sectorial investors.

Another aspect that is very relevant and that is mentioned in the Chapter 11 of the report: “Chapter 11: Weather and Climate Extreme Events in a Changing Climate”; is the increasing number of events, such as heat waves or storms, that were happening years apart. The spikes of these events can cause, among the great economic and physical damage, outlying occurrences in the data that can affect temperatures models and force a change in modelling to even more complexed and complicated methods (Diffenbaugh, N. S., Singh, D., Mankin, J. S., Horton, D. E., Swain, D. L., Touma, D., Charland, A., Liu, Y., Haugen, M., Tsiang, M., & Rajaratnam, B. 2017). An example that affects the matter of our research, which is the temperature in NY are the two snowstorms and cold spikes that happen in 2017 and 2018 (Blinder, A., Mazzei, P., & Bidgood, J. (2019, March 13))., in 2017 the weather touched the level of thirty degrees Celsius under zero.

Mathematic Stochastic Processes

A stochastic process, also known as a random process, is a mathematical concept used to model systems or phenomena that display random behavior over time or space. It involves a collection of random variables indexed by time or another parameter, with each random variable representing the value of the process at a specific point in time or space. Stochastic processes have widespread applications in fields such as biology, chemistry, ecology, neuroscience, physics, image processing, signal processing, control theory, information theory, computer science, cryptography, telecommunications, and finance (Liu B. 2010).

Stochastic processes provide a framework for understanding and analyzing the behavior of random phenomena. They allow us to capture the inherent randomness and uncertainty that exist in many real-world systems. For instance, a stochastic process can be employed to model the growth of a bacterial population, the fluctuations in an electrical current caused by thermal noise, or the movement of gas molecules. By studying the statistical properties of a stochastic process, such as its mean, variance, and higher-order moments, we can gain valuable insights into the dynamics and behavior of the underlying system.

There are various types of stochastic processes, each characterized by different mathematical properties and assumptions. Some commonly studied stochastic processes include the Wiener process or Brownian motion, which describes the random motion of particles in a fluid, and the Poisson process, which models the occurrence of events in continuous time. These processes play a fundamental role in the theory of stochastic processes and find numerous applications across different domains.

In summary, stochastic processes offer a mathematical framework for modeling and analyzing systems that exhibit random variation over time or space. They find extensive applications in diverse disciplines and have paved the way for the development of sophisticated mathematical tools and techniques to study random phenomena (Liu B. 2010). By utilizing stochastic processes, researchers and practitioners can gain a deeper understanding of the inherent randomness in nature and effectively model and analyze complex systems.

ARMA models

An autoregressive moving average (ARMA) model is a mathematical model commonly used in time series analysis to describe the behavior of a stochastic process. It combines autoregressive (AR) and moving average (MA) components to capture both the linear relationship between current observations and past values, as well as the influence of random shocks or errors. The AR component represents the linear dependence of the current observation on its own lagged values, while the MA component models the dependence on the random shocks or errors from previous time periods (Tsay, R. S., & Tiao, G. C. 1984).

The conditions for an ARMA model are that the data should be normally distributed and stationary, meaning that its statistical properties remain constant over time, and the autocorrelation function of the data should exhibit a decay pattern. Stationarity ensures that the model parameters can be estimated consistently, and the autocorrelation condition indicates that the current observation can be explained by a finite number of past observations.

ARMA models have various applications in finance. They are often employed to model and forecast financial time series data, such as stock prices, exchange rates, and asset returns. By capturing the temporal dependencies and random fluctuations in these data, ARMA models can assist in identifying patterns, making predictions, and estimating risk measures. Additionally, ARMA models can be extended to more sophisticated models like autoregressive integrated moving average (ARIMA) or generalized autoregressive conditional heteroskedasticity (GARCH) models, which are commonly used in asset pricing, risk management, and financial forecasting.

In summary, an ARMA model combines autoregressive and moving average components to describe the behavior of a stochastic process. It requires the data to be stationary and exhibit a decay in autocorrelation. In finance, ARMA models are widely used to analyze and forecast financial time series data, providing valuable insights for decision-making and risk management.

4. LITERATURE REVIEW

Weather Derivatives: A Risk Management Tool for Weather-sensitive Industries

The first literature review is about the paper that explains thoroughly how these derivatives work and is called: “Weather derivatives: A risk management tool for weather-sensitive industries”. It was published in April of 2000 by its authors Andreas Muller and Marcel Grandi. Two reinsurance finance experts based in Munich, Germany. The piece of literature is part of a larger volume called: “The Geneva papers on risk and insurance”, volume 25, second issue. Which covers many aspects in terms of risks and liabilities of the environment, nuclear power, and causing harm to the globe. It has twelve chapters, and they are written by all different authors, as a joint volume it takes different experts from all around and is considered still a valid piece of literature. Particularly, the chapter dedicated to weather derivatives is still adopted in many classes, for example, the international master in sustainable finance at Politecnico of Milano management school of Business cites this chapter and adopts it during the learning process in the course related to financial risk and sustainability rating.

It is key to denote how the authors thoroughly explain the main purpose of these options, how they function, and how they could be a great opportunity for companies that carry operational risk based on temperature or weather in general. These derivatives are very similar to XL/stop loss reinsurance contracts, as explained in the chapter, and are hedging assets. However, the chapter does raise a few doubts concerning the pricing model. As explained before, the HHD and CHD are the main factors for calculating the strike value of the derivative transaction, the way these

are calculated is by factoring a sum of the degrees under or over a predetermined value.
This paper is adopted in the thesis mainly for explaining purposes as it contains all the information required to understand how weather derivatives work and carries valuable credibility as research literature.

Market Expectations of a warming climate

The next piece of literature that will get revised is the paper called:" Market Expectations of a warming climate". This was carried out by researchers Wolfram Schlenker and Charles A. Taylor that consisted in picking 10 different locations in the US where there are historical weather daily data available and weather derivatives concerning the location. They used this information to compare prices of financial derivatives whose payouts were based on weather outcomes, these are the weather derivatives that are analyzed in the thesis, to the CMIP5 climate model predictions. This research showcases how the market directs reactions and measures when forecasting the weather to the constant increase in global warming over the last twenty years. To briefly sum they state how the pricing model incorporated and updated their forecast expectations during the climate events of the last twenty years.

This paper is of interest not mainly for the purpose, as it differentiates in terms of results as the authors focused mainly on the forecast level, but for the methodology of how the research was carried out. They took specific locations, in their case 10, and verified if they had the data availability in terms of weather and temperature and they determined them by taking the future prices of the weather derivatives traded at the Chicago Mercantile Exchange (CME). The historical data of the prices were retrieved on Bloomberg, which will be the same source for our research question.

Therefore, this literature paper is of great interest for their subject of research but mostly for how and where the data was retrieved, which makes the data valid the paper has been published in the journal of economics, volume 142 issue 2, and is available in many relevant locations, as Elsevier.

This paper is also the first paper that considers climate change expectations derived from weather-related future contracts. Furthermore, they show proof that financial markets do incorporate trends of future temperature model projections. This conclusion brings us to our larger-scale objective, which is providing evidence that companies' and industries' ability to create value is affected by future climate projections.

On Modelling and Pricing

The paper studied has the main objective to find a pricing model for weather derivatives with payouts depending on temperature. The research was carried out by Alaton, P., Djehiche, B. and D. Stillberger, D. in 2002. In the first section the research thoroughly explains how the derivatives work, in a similar way as it is here. The work that is the most of interest for the thesis is their temperature modeling they adopt a stochastic process, explained earlier, to explain temperature data, using historical data.

Upon analyzing the temperature data, they observe a strong seasonal variation. The mean temperature fluctuates between approximately 20°C during the summers and -5°C during the winters. Based on a preliminary examination of the data, they propose modeling the seasonal dependence using a sine function. This function takes the form *of* $\sin(\omega t + \varphi)$, where t represents time measured in days. Since the oscillations follow a yearly pattern, set $\omega = 2\pi/365$. To account for the fact that the minimum and maximum mean temperatures do not necessarily occur on January 1 and July 1, respectively, introduce a phase angle φ.

Additionally, a closer inspection reveals a weak positive trend in the data. This could be attributed to global warming or the urban heating effect, where temperatures tend to rise in areas near large cities. To capture this trend, they assume a linear approximation. While a polynomial trend could also be considered, the linear term dominates due to its minimal impact on the overall dynamics of the mean temperature.

Combining these considerations, it's constructed a deterministic model for the mean temperature at time t, denoted as TAVG (temperature average). The model takes the form:

$$TAVG = A + Bt + Csin(\omega t \ + \ \varphi) \quad (4)$$

where A, B, C, and φ are parameters to be determined.

However, temperatures are not deterministic but rather subject to inherent variability. To create a more realistic model, the researcher introduces a noise component to the deterministic model. Choosing a standard Wiener process, represented as $(Wt,\ t \geq 0)$, as the driving noise process. This choice is not only mathematically tractable but also aligns well with the observed daily temperature differences, which follow a normal distribution.

Furthermore, we notice that the quadratic variation of temperature, denoted as *σt^2*, varies across different months of the year but remains relatively constant within each month. Particularly during winter, the quadratic variation is significantly higher than during other months. Motivated by this observation, we assume σt to be a piecewise constant function, taking different values during each month.

To estimate the parameters A, B, C, φ, a, and σ, they utilize the temperature data from Bromma Airport over the past 40 years.
After estimating the parameters, the paper simulated the temperature and develops a complex pricing model for these options, they use two methods one of which is the Montecarlo simulation this is done by using the developed pricing and simulating multiple process trajectories and taking the arithmetic mean of the results to estimate the projected outcome. However, the paper does not compare the results to any real data so, as the paper is based on strong financial and mathematical knowledge, does not manage to give a perspective on how well it works.

Climate Change 2022 Mitigation of Climate Change

The climate change mitigation report of 2022 is the seventh report created by the IPCC; the organization body created by the united nation has the goal to provide scientific data about the environment useful for climate policymakers. The report is detailed and counts more than 1000 pages. The parts of our interest are mainly in the summary and specifically, the chapter of the summary called: "Recent Developments and Current Trends" and "Linkages between Mitigation, Adaptation, and Sustainable Development".
The chapters cover all the recent developments related to the environment and the consequence of continuous emissions in the atmosphere. It also deals with the future predictions of the environment and the consequence of the increase in temperature.

It is useful for the thesis as a source of information and for proving how the future, in terms of temperature and weather is uncertain. The risks these changes will carry are relevant for the companies and markets. As there are more than 80% of companies are affected by climate or weather conditions. The Climate change literature is the guideline, concerning the thesis, for what has happened and what is forecasted in terms of trends in climate change and emissions.

Climate Change and the Potential Use of Weather Derivatives to Hedge Vineyard Harvest Rainfall Risk in the Niagara Region

There is an interesting paper written by Don Cyr and Martin Kusy, professors in the department of finance at Brock University in Canada, on how potentially weather derivatives can be used to hedge against climate change that is affecting the agricultural industry, mostly Vineyards, in the Niagara region. The paper is called: "Climate Change and the Potential Use of Weather Derivatives to Hedge Vineyard Harvest Rainfall Risk in the Niagara Region"; was published in November 2010 and is present in the Journal of Wine Research. The subject of this paper is very useful, for the thesis, as it was written ten years ago and can be seen as a standpoint on how the purpose and the adoption of these derivatives have changed due to climate change.

In this case, the research question is if the change that is affecting the vineyard industry is possible to hedge with weather derivatives. The weather derivatives have been a plausible solution for vineyards, against extreme temperature events and amount of rainfall; however, there has been a slow rate of adoption by the actual participants of the industry. At the time is paper was written, there had been an important event that led to this paper which was the terrible harvest of the vineyards in California, which was a consequence of a season full of rainfalls that ruined the grape harvest. Therefore, as the Niagara region is known for its wine and grape liquors, such as Ice wine and Grappa, the paper shows how hedging can lower these risks. Specifically, as temperature options, the options that are discussed in the paper are priced based on a cumulative sum of the total rainfall during a period, which would ideally consist of the harvest season and not a season as summer or winter. This is an evolution from the basic option, and it consists of a more specific time, plus the weather data would be taken by pre-established weather stations of the region where the underlying (rain) of the option occurs. There are still some issues, such as a few days of unceasing rai, but an overall good season may disrupt the data regarding the underlying, as with the temperature pricing the cumulative sum does not protect against great fluctuations.

The paper also gives interesting insights, that will be considered while doing the methodology of research of the thesis, on how to approach and do analysis with weather conditions data and how to evaluate the pricing. In terms of estimating the rainfall as a variable, they adopt a times series analysis method plus an intervention analysis to prevent the presence of outliers (extreme events). Instead, to evaluate the options they adopt two methods: Burn rate analysis and Monte Carlo simulation.

5. DATA RETRIEVEMENT

Temperature data

The data is retrieved through the NOAA, the National Oceanic and Atmospheric Administration, specifically the National Climatic Data Center (NCDC) that is a department of the NOAA. This an organization responsible for understanding and predicting changes in Earth's environment, as well as conserving and managing coastal and marine resources. NOAA carries out various functions to fulfill its mission, which include creating and disseminating reliable assessments and predictions of weather, climate, the space environment, and oceanic and marine resources. They also produce nautical and geodetic products and services, implement integrated approaches to environmental management and ocean and coastal resources development, protect essential fish habitat, maintain sustainable fisheries, and work towards the recovery of endangered and threatened species of fish and marine mammals. Additionally, NOAA ensures access to sustained, reliable observations from satellites, ships, radars, and data buoys, and develops partnerships for the expansion and transfer of environmental knowledge and technologies.

One important aspect of the NCDC's work is the quality of its data. NCDC places great emphasis on ensuring the quality, objectivity, utility, and integrity of the information it disseminates. To achieve this, NOAA follows the OMB (Office of Management and Budget) and DOC (Department of Commerce) information quality guidelines. These guidelines provide policy and procedural guidance for federal agencies to ensure the quality of information they disseminate. NOAA implements these guidelines to ensure that the information it shares is accurate, reliable, and useful to the public. The data disseminated by NOAA and its departments undergoes quality control measures and is made freely available to the public through various data access tools. The data collections include oceanographic and geophysical data, underwater video, and imagery, supporting documentation, secondary products, and published works. The NCDC instead, offers all the data regarding temperature and weather related data of a location, such a daily temperature, rainfall or snowfall.

Furthermore, NOAA has implemented the NOAA Big Data Program, which aims to broaden access to its data resources. Through partnerships with Infrastructure-as-a-Service (IaaS) providers, NOAA and its departments leverages scalable computing capabilities to make its data more publicly accessible. The Big Data Program enhances full and open access to NOAA's data. Data made available through the program can be accessed for free and without restrictions on use. By leveraging cloud computing platforms, the program enables faster development of applications and analysis of NOAA's data, while reducing the load on NOAA's servers and networks. This program not only ensures the quality of data but also facilitates innovation and collaboration by making NOAA's data more accessible to a wider audience.

Bloomberg terminal

Bloomberg is a global provider of financial news and information, offering a wide range of services across its platform, television, radio, and print. It was founded by Michael Bloomberg in 1981 and initially focused on providing financial analytics and information. Over the years, Bloomberg has grown into a leading financial company with various offerings in the finance sector.

One of the key revenue earners for Bloomberg is the Bloomberg Terminal, also known as the Bloomberg Professional Service. This integrated platform provides real-time and historical price data, financials, news, and trading data to more than

325,000 customers worldwide. The Bloomberg Terminal is widely used by large institutional investors and offers a comprehensive array of financial capabilities. It allows financial professionals to track and analyze breaking news, communicate with other institutions, and access detailed information about financial markets globally. The terminal available to Bi student is the Bloomberg terminal used for the research of the CDD or HDD prices; however, to access data to all the HDD and CDD info available from the CME require an additional subscription not available to BI students. In fact, the optimal source for these derivatives would be the CME database or OptionX, which have incredibly high costs. Finally, after a thorough research there was data on CDD NY monthly contracts, which will be used for this research (Appendix 2).

The difficulty of finding real data on these weather derivatives is confirmed by the fact that there aren't any papers with actual data but mostly theoretic or experimental papers.
Bloomberg's influence and importance in the finance sector are substantial. It serves as a vital source of financial news and information for professionals in the industry. With 2,700 journalists and analysts in 120 locations, Bloomberg publishes over 5,000 news stories a day across multiple platforms, providing up-to-date and comprehensive coverage of markets, technology, politics, and opinion. This extensive network of information allows financial professionals to stay informed about market trends, economic indicators, and global developments that can impact their investment decisions.

Furthermore, Bloomberg offers various technological tools and solutions to its clients. These tools range from the Bloomberg Terminal, which caters to individual financial professionals, to flexible enterprise solutions that help integrate systems and automate workflows for financial market players. By providing these tools, Bloomberg assists its clients in making informed decisions based on reliable data and analysis.

Data management has become increasingly important in the finance sector, and Bloomberg recognizes this trend. The company offers data management services to financial firms of all sizes, helping them leverage datasets that provide insights into alternative information sources, ESG practices, regulatory compliance, supply chains, and more (Bloomberg Professional Services. (2021, October 7)). With the vast amount of data available today, financial companies need robust data management systems to handle the integration, cleaning, and analysis of data from multiple sources. Bloomberg's data management services assist companies in bringing together disparate datasets, making them available across the organization, and enabling better decision-making in areas such as research, trading, risk management, and accounting.

In summary, Bloomberg is a prominent provider of financial news, information, and data services. Its flagship offering, the Bloomberg Terminal, plays a crucial role in the finance sector by providing comprehensive financial capabilities and enabling professionals to stay updated on market developments.

5. METHODOLOGY

Fit data temperature

To start our process, we need to consider the steps made by the research of (Alaton, P., Djehiche, B. and D. Stillberger, D. 2002) before doing any pricing modeling for the options needed to adapt the temperature data in a way it was fit. As described by the literature review, they find at the end a daily average temperature (4) that is computed using complexed mathematical systems, from which they can forecast/simulate temperature trends that they later use in the pricing model developed and the Montecarlo simulation. These are the two models adopted to calculate a probable price for weather derivatives, due to a matter of replicability in this research the data temperature will be modelled similarly to the paper mentioned earlier (Alaton, P., Djehiche, B. and D. Stillberger, D. 2002), by applying these similar steps to the temperature data the research manages to properly simulate the temperature within our model.

For this research we will apply temperature from NY, from the JFK airport weather station, stationid: GHCND:USW00094789, since the only available option prices to BI University students on the Bloomberg terminal are CDD NY weather derivatives (Appendix 2). There are some differences in between the temperature threats that affect Stockholm weather, which was the one used in the research, but the model system base is the same, and using their mathematical findings and applying them in the code can help to identify the different seasonality and trends. However, there are some other differences that we can’t automatically solve, some assumptions are required before starting the process. The overall global increase in temperature could be similar but may have different consequences and may or may not be directly linked to global warming itself as big metropolitan cities are affected by many factors that impact the temperature. The uncertainties can negatively affect our simulation. Furthermore, recent events are showing that sudden temperature events are more often to happen due to the multiplier effect that global warming has on already existing phenomena that always took place (Diffenbaugh, N. S., Singh, D., Mankin, J. S., Horton, D. E., Swain, D. L., Touma, D., Charland, A., Liu, Y., Haugen, M., Tsiang, M., & Rajaratnam, B. 2017). Therefore, we are forced to assume a linear increase, as there is no alternative to include this increase into the model. This assumption is fundamental to respect the random nature of temperature but at the same time makes the findings a bit biased and unprepared towards future temperature behavior.

The first step in modeling the temperature is to retrieve the data, as aforementioned the source is the NCDC which gives the opportunity to download free daily temperature data, when available. As described before, the data arrives under a CSV file. The data downloaded corresponded to a daily time series, from “1970-04” to 2020-04”, of temperature. Here another methodological decision had to be made, as COVID in the last year and the quarantine have strongly affected the air quality, due to the stopping of factories and the emission of cars. The decision to not corrupt the data, as this COVID-related daily temperature data would not follow the yearly trend but is affected by a precise unique event.

After receiving the data by email the daily data temperature is imported using the read.csv call in R without changing or modifying the data. Here there is another important decision as the paper from Alaton, P., Djehiche, B. and D. Stillberger, D. uses the max temperature data. In this paper there were 7 years of temperature average available, from 2013 to 2020, and, since there are not papers around on options derivatives that use temperature average as a starting point, the decision was to opt for the temperature average as the basis of the simulation. This changes slightly in the plotting. Then the seasonality of the data is determined by regressing, using the lm R function, the dates of the temperature on the temperature itself which removes the linear trend present in the data, then we compute a periodogram that can indicate any dominant or cyclical periods. The next step in the model is to run the regression to find the deterministic trend of the data, using the formula (4) from the Alaton, P., Djehiche, B. and D. Stillberger, D. (2002) paper. Therefore, we first find the elapsed time and then design a matrix formula for the parameter's functions (4) in the R environment. Finally, find the deterministic trend by running the regression of the thesis where we regress the formula (4) onto the temperature data. Now that the necessary work is done and is possible to simulate the temperature, whereby using the seq R call dates are created. We only have interest for the short term, as we have the July 2021 NY CDD, but it could be interesting to see the forecast at a medium term. To forecast the temp, we adopt an ARMA model (arima (1,0,1) in R), with the conditions and assumptions imposed on the data is fitted for this prediction model. In the ARMA model, we use the residuals of the deterministic linear model. Following the ARMA model the trend model is adopted to predict (predict call in R) the deterministic trend in the simulated future time, then it's possible to simulate using the ARMA model (simulate call in R), after having the simulated residuals we add the predicted trend so that a TAVG is computed and not just a residual value. Through the imposed stochastic process, the temperature was forecasted and is fit for predicting the price of the NY CDD contracts.

5.2 How to Price CDDs

The pricing of CDDs as mentioned before is a very complicated matter, that has recently been discussed in more papers, such as (Ali, Z., Hussain, J. & Bano, Z. 2023) and (Alaton, P., Djehiche, B. and D. Stillberger, D. 2002); which are both very recent and are confirming the interest and curiosity for this case. For the thesis however it's not feasible to rely on complicated mathematical formulas as the ones suggested by (Ali, Z., Hussain, J. & Bano, Z. 2023) in their papers, as it would require high mathematical competencies and it wouldn't prove one of the reasons this thesis took place, which is the inconsistency between the way payoff and contract are stipulated and how difficult it is and will be in the future for the company to analyze and understand how to adopt these financial assets, because of all the temperature data, environmental, and economical related issues that are discussed earlier in the paper. Therefore, the pricing will be based on the payoff formula (3) and the ex-ante expectation CDD that is given by the simulation, as it would correspond to a minimum threshold that is forecasted before the option takes place; therefore, we have a fixed term CDD and an added variable that corresponds to the daily payoff expected when the daily temperature, if the time series differs from the determined temperature 65 degrees. The pricing adopted here is purely logical and based on the assumption of what a hedger would expect to pay (maximum) for the option, that is the expected total CDD plus the payoff, as we know by economic logic that anyone won't by purchasing something at more than its maximum value.

$Variable\ daily\ Payoff\ for\ 1\ tick = \max\,(K - \mathrm{C})$ (5)

The formula above is the payoff for the put option added to the predicted overall CDD. Therefore, as explained the price would consist of the payoff (variable per each temperature day) plus the basis CDD cost:

$Sim\ CDD\ = Variable\ Daily\ Payoff + E(CDD)$ (6)

The price calculation is fairly easy compared to the process used to model temperature and the ones that are present in the literature (Ali, Z., Hussain, J. & Bano, Z. 2023) and (Alaton, P., Djehiche, B. and D. Stillberger, D. 2002). However, the concept is straightforward and attempts a theoretical, easy to comprehend, approach using the starting concepts and information that are given when approaching the subject (CME)

The process above in R is done by doing a Montecarlo analysis, which is the most adopted model when forecasting weather options (Ali, Z., Hussain, J. & Bano, Z. 2023) (Alaton, P., Djehiche, B. and D. Stillberger, D. 2002) (Don Cyr, Martin Kusy, Anthony B. Shaw (2010). The Montecarlo simulation inputs 1000 different simulated temperatures, from the trend model already built.

Least squared Errors

Following the Montecarlo analysis the quality of our model can be tested by doing the LSE of the average CDD price with the actual prices for the monthly CDD price for July in New York.

$LSE = \sum_{1}^{t}(actual\ CDD - sim\ CDD)^2$ (7)

In R we simply input the formula and run it.

Minimizing parameters

Finally, the question is how this research can contribute on both an empirical and theoretical level offering an added piece on how we can better perform complex temperature models, to price better weather derivatives.

This contribution is in the form of a parameter alfa to solve the function of the LSE, keeping as a variable the simulated pricing and how we can improve it.

$MinLSE = \alpha * LSE$ (7)

In R we adopt the optimise call, having as variable input the simulated CDD prices.

6. RESULTS AND INTERPRETATION

Temperature data

The first step is importing the data, subset it to the dates and temperature type that we have interest in and plot it.

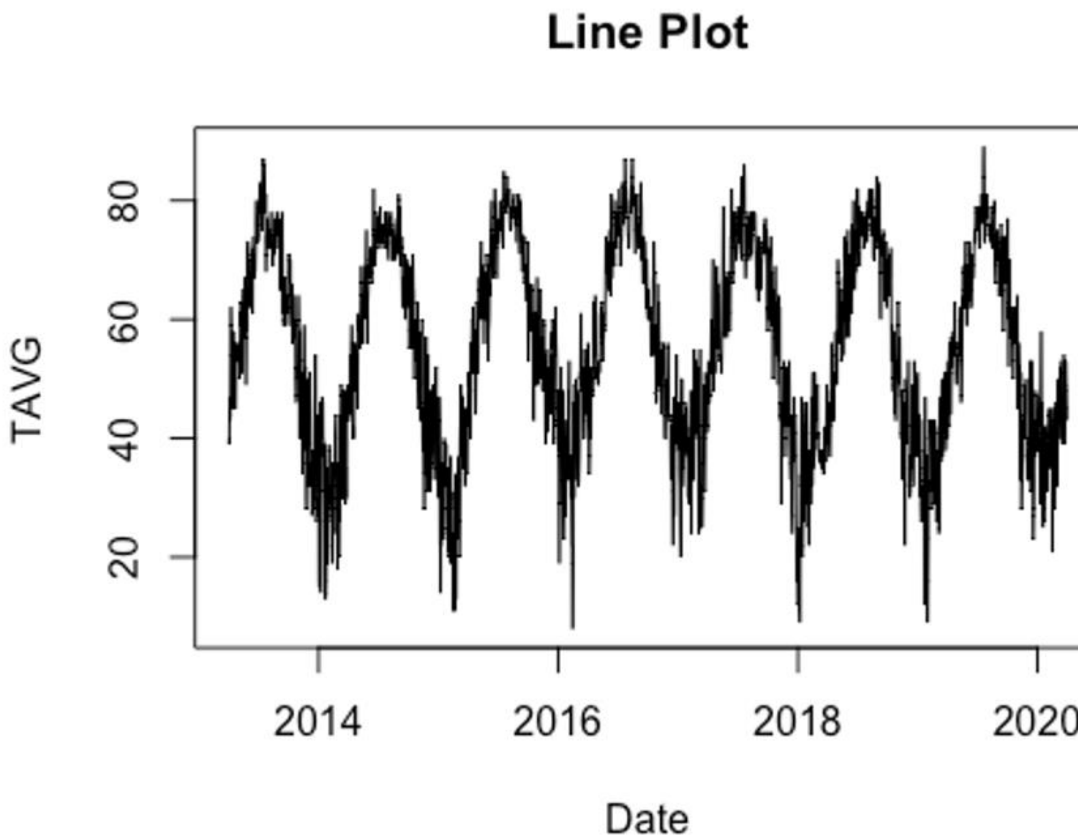


Figure 3: Unmodelled Temperature trend from 2013 to 2020

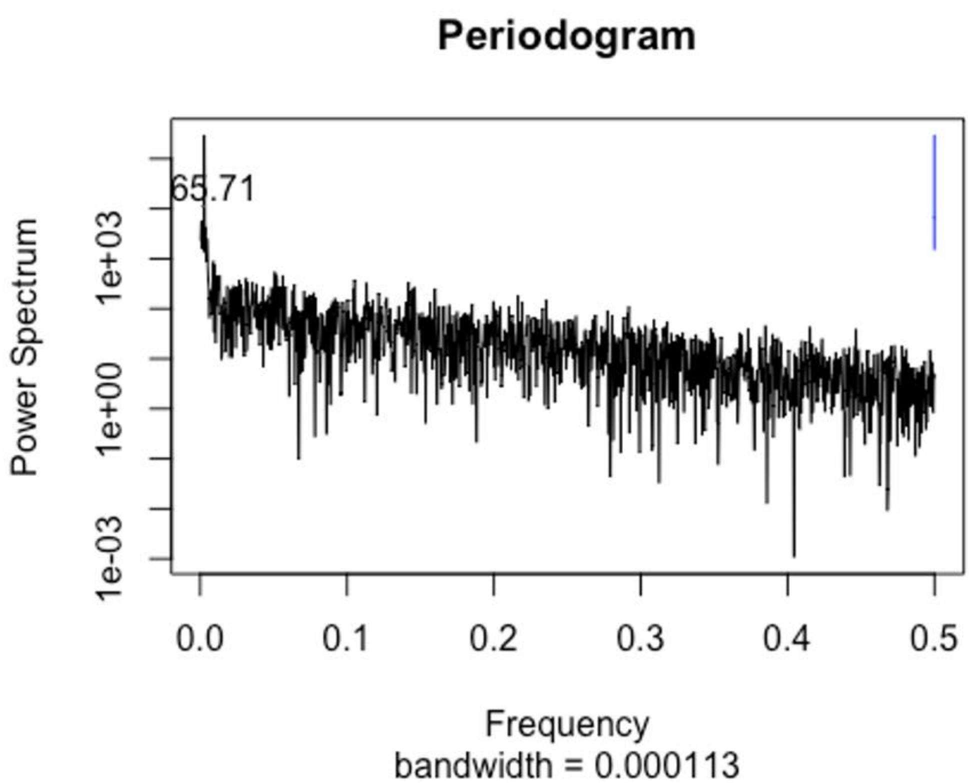
As it’s possible to see from the plot, is already clear how the two cold spikes and snowstorms of late 2017 and late 2018 impacted the trend, we expect to have different results when compared to the research temperature of Stockholm from the Alaton, P., Djehiche, B. and D. Stillberger, D. paper.

Modeling temperature

The second step is to work on the data, as explained in the methodology, in a way that we can get temperature data that can be utilized in the forecasting model. From removing the linear trend through finding the residuals of the first regression:

```
TAVG2013$Detrended <- residuals(lm(TAVG2013$TAVG ~ TAVG2013$DATE + 0))
```

From this we get the following periodogram:




```
print (topFrequencies)
## [1] 0.002734375 0.001953125
print (topPeriods)
## [1] 365.7143 512.0000
```

Figure 4: Periodogram

The periodogram shows a slightly decreasing frequency trend. Therefore, we need to assume that these events will linearly continue to happen, which will cause a decrease in the trend of the temperature. Top periods show when these events take place: 364.71 and 512. And the top frequencies: 0.002734 and 0.00195. It's feasible to state how there is a repeating trend based on the time of the year, as expected.

The second step of the process is to fit the deterministic trend into the TAVG parameters formula and run (4), from which we get the following trend model regression:

```
Call:
lm(formula = TAVG2013$TAVG ~ designMatrix(elapsedTime))

Residuals:
     Min       1Q   Median       3Q      Max 
-26.3096  -3.9461  -0.0279   4.1423  24.9095 

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)    
(Intercept)                54.45701    0.26162 208.153  < 2e-16 ***
designMatrix(elapsedTime)1  0.23548    0.06483   3.632 0.000286 ***
designMatrix(elapsedTime)2 -9.72608    0.18415 -52.817  < 2e-16 ***
designMatrix(elapsedTime)3 19.50138    0.18537 105.202  < 2e-16 ***
designMatrix(elapsedTime)4 -0.02627    0.18415  -0.143 0.886571    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.584 on 2551 degrees of freedom
Multiple R-squared:  0.8452,	Adjusted R-squared:  0.8449 
F-statistic:  3481 on 4 and 2551 DF,  p-value: < 2.2e-16
```

Table 2: Summary TAVG (4) regression

Where design Metrix (4) is defined as:

```
designMatrix <- function(t) {
 matrix(c(t, cos(2 * pi * t), sin(2 * pi * t), cos(4 * pi * t)), ncol = 4)
}
```

To interpret the regression, we start by analyzing the significance level of the coefficients, it appears that the intercept, which is always kept even if not significant, and the first three parameters have very significant as their P-value is close to zero. The fourth variable is not significant at any level, this was threatening at first. However, after a brief research, a paper published on the 26th of May showcases a modified version of (4) with the three parameters significant plus a new parameter D (Ali, Z., Hussain, J. & Bano, Z. 2023):

$TAVG = A + Bt + Csin(\omega t) + Dcos(\omega t)$ (8)

This confirms the quality of the regression, and it encourages to continue with the model. The R squared is fairly high, this is not a very good sign as it can prove the presence of multicollinearity, which means that the variables may not be fully independent, and the results may be biased. This is also portrayed in the Alaton, P., Djehiche, B. and D. Stillberger paper and is expected when working with temperature data.

After running the regression, we can add the trend that was found earlier and plot the modeled deterministic trend (figure 5) that will be utilized to simulate the future TAVG.

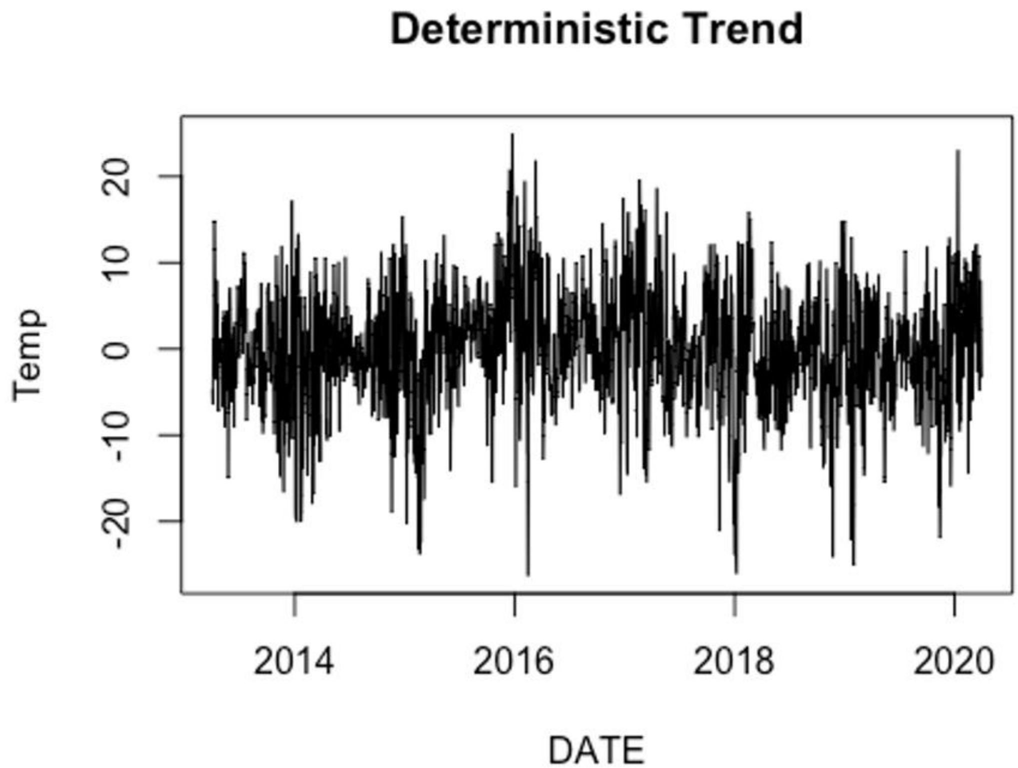


Figure 5: Deterministic trend TAVG 2013-2020

From looking at the plot we can see a very slight increase in temperature, the causes are unknown as it can be a consequence of global warming (figure 2), of an increase of gas emitted from the city, the influence of energy consumption; most likely is a mix of all the above plus other factors. If we compare it with the Stockholm temperature model presented in the paper, there is confirmation that in NY there were extreme events, and they are predicted to happen again in the future. We can state that global warming does affect the modeling, as we had included in the deterministic trend, and it's possible to only assume its effects in the future, from affecting a temperature model to extreme events, as there is no model that is able to predict such events.

Simulating temperature

The initial step, as stated before, is to fit the ARMA model. Using as simulating regressor the trend model residuals. When predicting the temperature there are long time series with the predicted temperature, to which we add the calculated deterministic trend (figure 6).

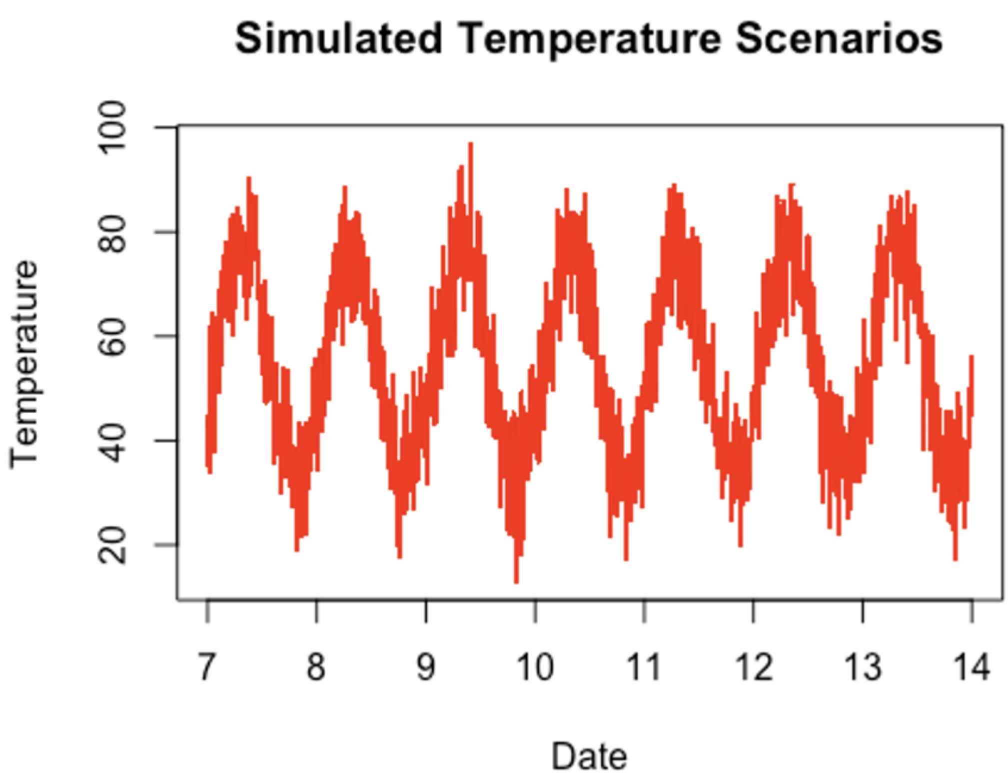


Figure 6: Simulated temperature years 2020-2027

The easiest and most straightforward way to interpret the simulation is to analyze the plot, its observable that in the following 7 years the temperature is expected to continue increasing, with two extreme temperature events, one in the summer of 2022 with a heat wave, and one in the winter of the same year with a cold spike. These forecasts are purely mathematical and do not have any reliability in terms of such events, as if they had they would be very precious and frequently used to avoid catastrophes.

If the temperature continues to increase, in an exponential way, the payoff formula calculation may have to change, in a way that an option holder doesn't lose or gain capital due to the perpetual increase in temperature. When analyzing how to face it, it's possible to compare it to inflation, as the economic system adjusts to inflation there could be a way that these options may be interconnected with the temperature increase so that there isn't any systematic error.

Montecarlo simulation as pricing

Before starting the Montecarlo simulation, the model verifies the number of CDD days in the simulation, calculated through (2) to check whether the research can proceed to the pricing (figure 7).

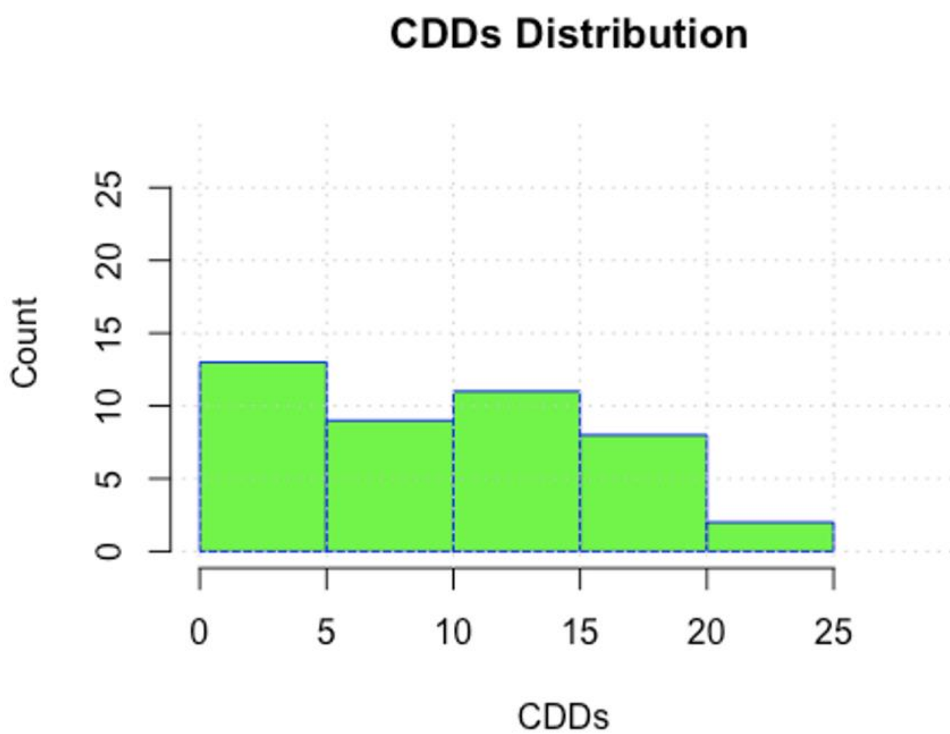


Figure 7: simulated CDDs density distribution

The histogram confirms the value of the model as there are more CDDs that are closer to the threshold number and a few CDDs far from 65 degrees. If the plot had a broader perspective in terms of time, for example, the full season, or a year it would have been possible to see a Gaussian distributed histogram.

The model then computes the simulated CDDs for a single repetition to establish the pricing and finds CDDs=$399,65, which is the minimum expected payment (that changes for each simulation) to which the pricing adds a variable scenario when a hedger would execute the put option and its payoff (3) on a specific day. After doing such a process within the Montecarlo simulation for n=1000 times the results are 1000 different CDD prices time series, for the month of July, and it's possible to plot the MC pricing distribution (Figure 8).

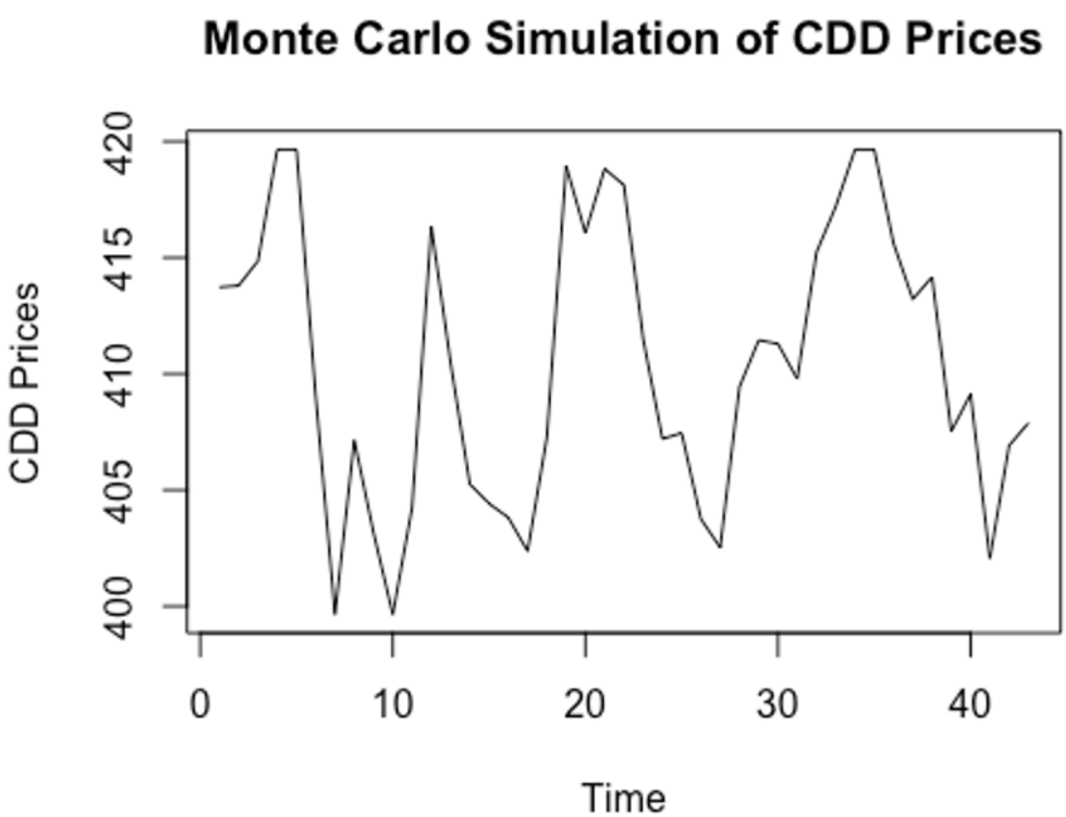
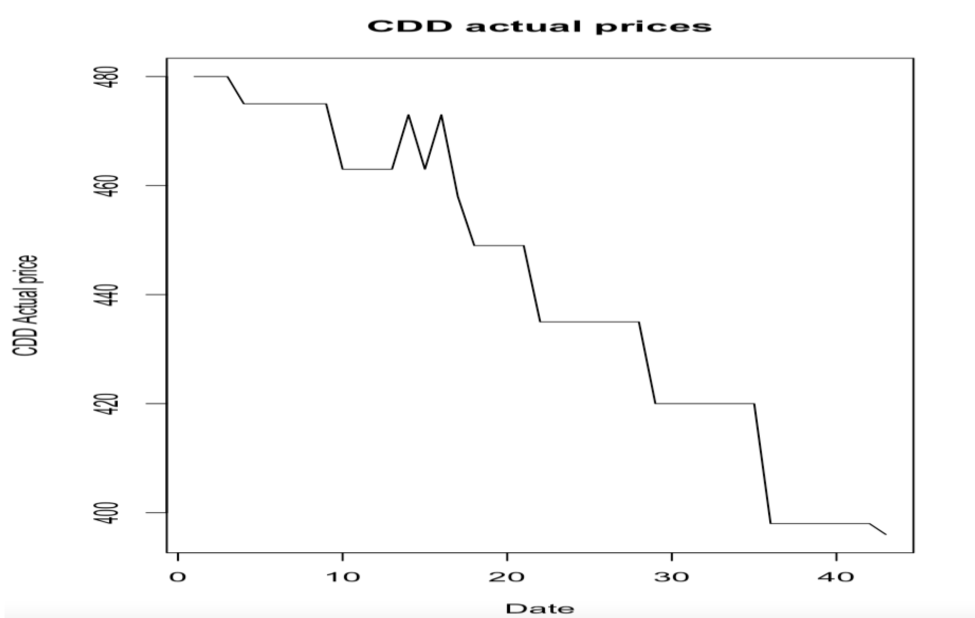


Figure 8: Monte Carlo Simulation of CDD prices

From reading the plot there are not any surprises as we expect to have a minimum variable, threshold, but that is similar in all n times as it represents the predicting model expectations on the temperature data, which in the computed pricing would represent what a possible investor can expect to pay when adopting a strategy involving NY CDDs July contracts. There is a big issue with this, as the model does not really price the option, but it values it giving it a value based on a single tick and based on the concept that we expect a min price when in a fact when an option is not in the money is not logical to execute and therefore would have value 0. The thesis is aware of this but takes the decision to have a condition where when there is no payoff there is still a value for the contract.

LSE

The final part of the model is the inclusion of the actual CDD prices, which makes the model quite unique as there is an attempt to have an empirical proof for this quite theoretical form of pricing. By importing the actual NY CDD prices, retrieved from Bloomberg, and starts analyzing the differences through graphical representation (figure 9).



It’s clear how the nature of a price and its maturity affects greatly the trend, as it starts at a higher price when there is no knowledge of the summer temperature, so people are hedging more, as times pass there is more information and prices drop. The positive side is that it doesn’t go below a certain price, around 380, which is shared with the model. This gives confidence that the model is on the right track, but there is still a lot of work to do and there are excitant mathematical formulas that are seen in this research in the literature that perform better.

The final contribution is a min parameter that can be used as another parameter for the simulation tool to minimize the error for this pricing method, the idea is replicable for any future and already adopted pricing, as it is a minimizing function on R applied to the LSE function (7) and we found the parameter, denominated alfa, equal to:

$$\alpha = 99.97$$

That consists of the final contribution of the model to the existent research on weather derivatives.

8.CONCLUSIONS

Before carrying out the final conclusions, is important to add a business perspective, as read in this paper the research available on weather derivatives is very theoretical, with notions and competencies that require a very specific set of skills. However, the concept of these financial assets is very practical, it's an insurance for phenomena that affect almost every company in the world, mostly small and medium enterprises that do not have the capital to face physical, transition, and operational risks related to weather and the effects of global warming. To make these options mainstream there is a clear necessity to have a standard and well recognized pricing model, starting from the modeling of temperature. This thesis offers a way to value these options, from temperature modeling to a value model. Both need more implementation as the pricing does not account for interests and market risks, which would add important details to the evaluation.

On the academic level, this paper offers an additional perspective: the inclusion of real data and a minimizing parameter. The parameter needs further inspection and will be fundamental to understanding its relationship with the simulated temperature. To conclude, the thesis question was "*How to price weather derivatives, and how the temperature of these can be modeled to be fit for forecasting? What are the effects of global warming on these?"*. The thesis managed to create a temperature model fitted for simulation, going through a stochastic process, and a pricing model that gives an idea of what the daily pricing of these options may be. Throughout the thesis there was a special focus on global warming and how is affecting temperature modeling and pricing. It is feasible to state that if society is not able to solve this major issue, its effects would go beyond modeling and pricing of weather derivatives.

REFERENCES

Campbell, S. and Diebold, F.X. Weather forecasting for weather derivatives, Journal of the American Statistical Association, 100, 6-16 (2005)

Tesfahun Berhane, Aemiro Shibabaw, Gurju Awgichew & Assaye Walelgn (2020) "Pricing of weather derivatives based on temperature by obtaining market risk factor from historical data"

Wolfram Schlenker, Charles A. Taylor (2021) "Market expectations of a warming climate"

Irene Monasterolo, Stefano Battiston, Yannis Dafermos (2021)."Climate Risk and Financial Stability".

Don Cyr, Martin Kusy, Anthony B. Shaw (2010). "Climate Change and the Potential Use of Weather Derivatives to Hedge Vineyard Harvest Rainfall Risk in the Niagara Region".

Andreas Muller, Marcel Grandi (2000). "Weather Derivatives: A Risk Management Tool for Weather-sensitive Industries".

"Climate Change 2022 Mitigation of Climate Change", IPCC 7th Climate Change report (2022).

Alaton, P., Djehiche, B. and D. Stillberger, D. (2002) On Modelling and Pricing Weather Derivatives. Applied Mathematical Finance. 9(1): 1–20, 2002.

Ali, Z., Hussain, J. & Bano, Z. (2023). Pricing weather derivatives in an uncertain environment. Nonlinear Engineering, 12(1), 20220291.

Dai L, Fu Z, Huang Z. Option pricing formulas for uncertain financial market based on the exponential Ornstein-Uhlenbeck model. J Intell Manuf. 2017;28(3):597–604.

Liu B. Uncertainty theory: A branch of mathematics for modeling human uncertainty. Berlin Heidelberg, Germany:Springer; 2010
Weather Options Overview – CME Group.
https://www.cmegroup.com/education/articles-and-reports/weather-options-overview.html

Tsay, R. S., & Tiao, G. C. (1984). Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Function for Stationary and Nonstationary ARMA Models. Journal of the American Statistical Association, 79(385), 84–96.

Diffenbaugh, N. S., Singh, D., Mankin, J. S., Horton, D. E., Swain, D. L., Touma, D., Charland, A., Liu, Y., Haugen, M., Tsiang, M., & Rajaratnam, B. (2017). Quantifying the influence of global warming on unprecedented extreme climate events. Proceedings of the National Academy of Sciences of the United States of America, 114(19), 4881–4886.

Blinder, A., Mazzei, P., & Bidgood, J. (2019, March 13). 'Bomb Cyclone': Snow and Bitter Cold Blast the Northeast. The New York Times.
https://www.nytimes.com/2018/01/04/us/winter-snow-bomb-cyclone.html

National Centers for Environmental Information (NCEI). (n.d.). Order History | Climate Data Online (CDO) | National Climatic Data Center (NCDC). Climate Data Online (CDO). https://www.ncdc.noaa.gov/cdo-web/orders?id=3367577&email=nico.algie18@gmail.com

(Bloomberg L.P.). Terminal. BI subscription.

Bloomberg Professional Services. (2021, October 7). The increasing importance of data management for financial firms of all sizes | Insights | Bloomberg Professional Services. https://www.bloomberg.com/professional/blog/the-increasing-importance-of-data-management-for-financial-firms-of-all-sizes/