CATÓLICA LISBON BUSINESS & ECONOMICS

Forecasting Flight Delays with Climate Data and Implications for the Airline Industry

Anna Christina Wimmer

Dissertation written under the supervision of professor Miguel Nogueira

Dissertation submitted in partial fulfilment of requirements for the MSc in Business Analytics, at the Universidade Católica Portuguesa, 02.01.2024.

Abstract

This thesis investigated the impact of climate change on the number of flight delays and corresponding flight delay costs by examining departure delays from John F. Kennedy Airport in New York City, USA, from 2013 to 2022 and deriving a model for future delays from 2023 to 2030. This is relevant as the airline industry faces high flight delay costs caused by a changing and more extreme climate. First, machine learning algorithms were trained and evaluated on past weather and flight delay data to find the best model to predict whether a flight is delayed or not and the cost of the delay. The best-performing models, a gradient boosting classifier and a gradient boosting regressor, were then used to make predictions on data of two future climate scenarios. These scenarios represent the upper and lower thresholds of the expected evolution of anthropogenic greenhouse gas emissions and resulting climate change. The outcomes showed no significant change in the number of weather-related flight delays and flight delay costs until 2030 based on the computed Kendall Taus and Spearman Rank Correlations. Additionally, the results identified significant differences in the average delay cost per flight between airlines. It was recommended to regularly repeat this research to spot increasing delay risks as early as possible. This thesis applied the business analytics principles by exploring how the airline industry can use the prediction results to make business decisions. Suggestions were cost reduction measures and increasing the quantity or prices of plane tickets or complementary services.

Title: Forecasting Flight Delays with Climate Data and Implications for the Airline Industry

Author: Anna Christina Wimmer

Keywords: Flight Delays, Flight Delay Cost, Prediction, Weather Data, Climate Change, Machine Learning

Resumo

O sector das companhias aéreas enfrenta custos elevados devidos a atrasos de voos causados por um clima em mudança e com cada vez mais eventos extremos. Por conseguinte, esta tese investigou o impacto das alterações climáticas no número de atrasos de voos e nos custos dos mesmos. O foco foram os atrasos nas partidas do Aeroporto John F. Kennedy, na cidade de Nova Iorque, durante um período de 2023 a 2030. Em primeiro lugar, os algoritmos de aprendizagem automática foram treinadas e avaliados numa base de dados histórica contendo dados meteorológicos e de atrasos de voos com o objetivo de encontrar o melhor modelo para prever se um voo está ou não atrasado e os custos associados. Os modelos com melhor desempenho, um gradient boosting classifier e um gradient boosting regressor, foram depois utilizados para fazer previsões sobre dados de dois cenários climáticos futuros. Estes cenários representam o melhor e pior cenários de evolução das emissões com origem antropogénica e as alterações climáticas associadas. Os resultados não revelaram alterações significativas no número de atrasos de voos relacionados com as condições meteorológicas e nos custos dos atrasos de voos até 2030. No entanto, indicaram o desenvolvimento de mais atrasos e custos mais elevados no futuro. Por conseguinte, foi recomendada a repetição regular desta investigação para detetar o aumento dos custos o mais cedo possível. Esta tese aplicou os princípios de business analytics, explorando a forma como o sector das companhias aéreas pode utilizar os resultados das previsões para tomar decisões empresariais.

Título: Previsão de atrasos de voos com dados climáticos e implicações para o sector das companhias aéreas

Autor: Anna Christina Wimmer

Palavras-chave: Atrasos nos voos, Custo dos atrasos nos voos, Previsão, Dados meteorológicos, Alterações climáticas, Aprendizagem automática

Table of Contents

List of Figures

List of Abbreviations

FAA	Federal Aviation Administration
JFK Airport	John F. Kennedy Airport
CMIP	Coupled Model Intercomparison Project
ScenarioMIP	Scenario Model Intercomparison Project
SSP	Shared Socioeconomic Pathway
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA5	ECMWF Reanalysis v5
MAE	Mean Absolute Error
MSE	Mean Squared Error
QDM	Quantile Delta Mapping

1 Introduction

The Federal Aviation Administration (FAA) reported a total cost of $33 billion for flight delays in the USA in 2019 with an increasing trend. These include costs for the US airline industry of $8.3 billion which have a growing impact on the industry's operating revenue of $196.2 billion (Federal Aviation Administration, 2020; Bureau of Transportation Statistics, 2020). Climate risk plays a significant role in causing flight delay costs. Climate risk is defined as a combination of hazard, exposure, and vulnerability (Fakhruddin, Boylan, Wild, & Robertson, 2020). In this dissertation the hazard for the airline industry is represented by extreme weather conditions. The exposure to changing climate and weather conditions is portrayed by the flight schedule of John F. Kennedy Airport (JFK Airport), as complications in flying, take off, and landing arise (Koetse & Rietveld, 2009). Furthermore, vulnerability comprises the costs resulting from weather induced flight delays to airline companies flying from JFK Airport.

Existing research proves that the climate is expected to change further, and the number of flight delays might therefore rise as well. The extent to which the climate will change is uncertain, dependent on reaching political climate goals among other factors, and can be highly heterogeneous across different regions (Pörtner, et al., 2022). This uncertainty has led to the development of numerical models of the earth's climate system which allow to better understand its sensitivity to anthropogenic greenhouse gas emissions and project its evolution under different future climate scenarios (Tebaldi, et al., 2020).

Several recent studies have investigated the impact of weather on flight delays or predicted flight delays using data science (L. Carvalho, 2020). However, the following questions remain open and will be addressed in this thesis:

- How will the number of weather-related flight delays develop from 2023 to 2030 compared to the past and present based on different climate scenarios?
- How will the cost of weather-related flight delays develop from 2023 to 2030 compared to the past and present based on different climate scenarios?
- What business implications can be drawn from forecasting flight delays and their costs for the airline industry?

This research is organized as follows: First, the previous research in the areas of climate risk, flight delays, and business impact of flight delays and their cost as well as the connection between these areas was examined. This allowed to build on past findings and helped to determine the right empirical methods and the required data for this research. Second, the data chapter described the used datasets on flight delays, climate data, and flight delay cost data. It included a detailed description of how the datasets were obtained, cleaned, merged, and prepared for modeling as well as descriptive statistics of relevant variables. Third, the methods chapter disclosed the training and evaluation of several machine learning models to select the models with the highest accuracy. It explained how the best-performing model made predictions of flight delays and costs for the years from 2023 to 2030. The results chapter described the findings of the methods mentioned before, including model performance and prediction results. The discussion chapter explored how airline businesses can deal with the costs of flight delays from a manager's point of view. Finally, this thesis included a section on the research limitations and the conclusion. This research investigated departure delays from JFK Airport exclusively and predicted values for a specific period (2023 - 2030). The conclusion summarized the most important findings and highlighted implications for the airline industry combining the disciplines of climate science and business analytics by investigating how future climate conditions will impact flight delay costs for airlines.

2 Literature Review

2.1 Climate Risk

2.1.1 Climate Change as the Cause of Extreme Weather

There has been abundant research on climate change in the past decades. The change in climate due to human emissions of greenhouse gases is now a well-established fact (Pörtner, et al., 2022). Greenhouse gases absorb the infrared radiation emitted by the earth's surface leading to a rise in the planet's temperature. As human activities lead to an unnatural increase in greenhouse gases, the temperature on earth increases more than it would naturally (Slama, 2016). This has led to dangerous developments, e.g., the cause of extreme weather events (Hertzberg, Siddons, & Schreuder, 2017). Examples of greenhouse gases include carbon dioxide which results from activities such as burning fossil fuels and deforestation, methane resulting from agricultural activities and transportation, and fluorinated gases commonly used in refrigeration systems, foaming agents, and fire extinguishers amongst other applications (Sovacool, Griffiths, Kim, & Bazilian, 2021).

Climate change can cause broad variations in weather patterns, which in turn can impact the environment, the economy, and society in diverse aspects. One example is more extreme precipitation events, resulting from the warming oceans (Gössling, Neger, Steiger, & Bell, 2023). This effect depends on how close an area is to water and mountains as well as on airflow direction. Furthermore, precipitation is expected to increase in areas closer to the polar regions, while it will decrease closer to the tropic regions (Koetse & Rietveld, 2009). In turn, a decrease in rainfall can lead to severe droughts and increase the frequency of heatwaves (Gössling, Neger, Steiger, & Bell, 2023). Additionally, climate change impacts wind speed and the frequency and intensity of storms (Borsky & Unterberger, 2019). Another consequence of climate change is a rising sea level which can lead to extensive coastal floodings and the destruction of infrastructure (Frederikse, et al., 2020).

In conclusion, previous studies have extensively investigated and demonstrated the changes in variety, frequency, and consequences of extreme weather events associated with anthropogenic emission of greenhouse gases. It is important to highlight differences in the type, magnitude, and future changes of extreme weather events across different regions and seasons (Koetse & Rietveld, 2009).

2.1.2 Climate Modeling and Data

For a long time, it has been the goal of climate scientists to study climate change and make projections for the future climate. Climate modeling emerged as a discipline within climate science to simulate climate dynamics, e.g., the earth's capacity to retain heat, patterns of energy storage and distribution, and circulation (Edwards, 2011). These climate models consist of mathematical equations that describe temperature, winds, ocean currents, and other variables related to climate. In present times a large number of climate models exist that are constantly compared and evaluated. Research facilities and other functional organizations across the globe are cooperating with the common goal of establishing the best unified climate model with the most accurate predictions of future climate (Edwards, 2011).

Several climate models have proven their accuracy and established their status over the years. One example of successful modeling is the Coupled Model Intercomparison Project (CMIP) which is currently in phase 6 (Tebaldi et al., 2020) providing a wide range of historical and future climate projections simulated with a wide range of state-of-the-art numerical climate models. CMIP6 models are a cornerstone of the Intergovernmental Panel on Climate Change reports, and a key source of information about the potential consequences of climate change, supporting key stakeholders and policymakers in making decisions to mitigate and adapt to future climate changes (O'Neill, et al., 2016).

CMIP6 includes the Scenario Model Intercomparison Project (ScenarioMIP) and models past as well as future climate scenarios based on climate development. ScenarioMIP consists of eight scenarios for the 21st century and several long-term extensions. Each of these scenarios is based on different Shared Socioeconomic Pathways (SSPs) representing a wide range of different plausible future evolutions, emissions, and changes in land use. Tier 1 scenarios represent a broad variety of future forcing pathways, from very high to very low emissions, and are considered to be of higher priority. They include, e.g., SSP5-8.5, SSP2-4.5 and SSP1-2.6. Tier-2 scenarios explore complementing scientific and political issues and include, e.g., SSP4-6.0, SSP4-3.4, and SSP5-3.4-OS (O'Neill, et al., 2016).

Another example for data often used in climate science is the ECMWF Reanalysis v5 (ERA5) data by the European Centre for Medium-Range Weather Forecasts (ECMWF). ERA5 is an atmospheric reanalysis model providing detailed global climate data by combining mathematical models with past climate observations. This leads to a comprehensive representation of past climate. It contains hourly data, high spatial resolution, and a component to estimate uncertainties. The model can be used for various climate research applications, e.g., it is often used for measuring wind power (Hersbach, et al., 2020).

2.2 Flight Delays

2.2.1 Extreme Weather as a Cause of Flight Delays

A large amount of literature exists about the impact of extreme weather on flight delays as a result of climate change, demonstrating that various weather events can affect flight delays in different ways. In these studies, flight delays are generally categorized into three types: in-flight delays, arrival delays, and departure delays.

In-flight delays can be caused by lighting, wind, precipitation, fog, icing and other turbulence. Lightning strikes display an extreme case as they could damage the airplane's surface and jeopardize the flight. The same is true for strong turbulence and wind which may cause discomfort and injuries for passengers and the crew. Fog and precipitation affect the pilot's visibility and heavy precipitation can restrict the pilot from overflying thunderstorms. Furthermore, strong wind can influence maneuverability and thereby increase fuel consumption and flight time. As a result, in-flight disruptions may lead to the decision to change the route, e.g., to evade thunderstorms, and hence, lengthen flight time (Škultétya, Jarošová, & Rostáš, 2021).

Arrival and departure delays are affected by similar weather factors. Various factors can affect the time airplanes spend on runways, including precipitation, wind speed, wind direction, ice, snow, hail, and heat (Lui , Hon, & Liem, 2022). As a result, either take-off distances need to be longer or, if not possible, the payload capacity needs to be reduced (Gratton, Williams, Padhra, & Rapsomanikis). Moreover, strong winds at airports can increase the risk of accidents and often impede airplanes from landing or departing (Koetse & Rietveld, 2009). Another long-term risk for airports in the vicinity of the sea or inland waterways is floods as a result of extreme precipitation. In 2018 Christodoulou et al. estimated a total of 196 airports to face inundation risk by 2080. They also found that increasing wind speeds can damage airport infrastructure and cause flight delays or cancellations (Christodoulou & Demirel , 2018).

As the different flight delay types show differences in their cause, extent, and implications, most papers tend to focus on only one delay type. The focus for this thesis was chosen to be departure delays after evaluating the available data sources. Therefore, predictor variables recommended for this research were variables related to precipitation, wind, snowfall, hail, or extreme temperature.

2.2.2 Flight Delay Data Analysis and Predictive Models

In order to choose the best approach for the data exploration and prediction of this thesis, relevant methods of past literature were evaluated. In general, it can be deferred between descriptive, predictive, and prescriptive types of analysis. While descriptive analysis means inspecting the relationships in past data, predictive is about forecasting future values, and prescriptive derives recommendations based on the predictive outcomes informing real-world decisions (Roy, Srivastava, Jat, & Karaca, 2022).

Each analysis type includes several methods. The descriptive type mainly contains statistical methods, such as correlation analysis, multivariate analysis and differences-in-differences. The goal of a descriptive analysis is to understand why certain patterns occur in the data and what variables play a role in specific outcomes (Roy, Srivastava, Jat, & Karaca, 2022).

Examples of methods applied in predictive studies are regression, network analysis, time series analysis, and machine learning (Mishra & Silakari, 2012). One example of past predictive studies on flight delays was conducted by Kim et al. in 2016. In their research, they used a deep-learning recurrent neural network to predict the flight delay status per day using on-time performance and weather data. The model achieved high accuracy using daily average values of the weather data. This proves that daily values provide significant information for predicting flight delays (Kim, Choi, Briceno, & Mavris, 2016).

Another example of existing literature on the prediction of weather induced flight delays is a paper by Mokhtarimousavi et al. They used a random parameter logit model as well as a support vector machine model trained by the Artificial Bee Colony algorithm to predict whether a flight is delayed or on time. The training of the models is based on arrival data over three months at Miami International Airport. Important insights on input variables can be gained from their research and will be essential when choosing the input variables for this thesis as well. The findings show that the date, the departure time, whether a flight is on a weekday or weekend, and the operating carrier deliver important information for flight delay prediction. Additionally, the flight delay cause, including whether the delay was due to weather conditions, was used as a predictor (Mokhtarimousavi & Mehrabi, 2023).

These examples illustrate the existing variety of analyses and forecasting methods on the topic of flight delays in scientific literature. These previous studies have mostly focused on the prediction of flight delays. However, there is a lack of prescriptive studies to draw real business conclusions. While prescriptive analytics methods are relatively novel in the context of flight delay analysis, one application would be analyzing real-time data to support ad-hoc decision-making (Frazzetto, Nielsen, & Šikšnys, 2019). This is an important novel contribution of this thesis, providing a framework for assessing the impact of weather induced flight delays for airline businesses, and its future evolution under a changing climate.

2.3 Business Impact of Flight Delays

2.3.1 Measuring the Business Impact of Flight Delays

A key focus of this research was business implications for the airline industry. There is limited literature available that quantifies the cost incurred due to a flight being delayed. One exception is provided by Airlines for America, a cooperation of American airline businesses to shape policy and promote safety and other industry interests. The association regularly publishes industry data, including statistics on flight delays which will be used in the analysis of this thesis. Specifically, they found an average delay cost of U.S. passenger airlines of $101.18 dollars per minute. Fuel expense and labour were the highest cost factors with $42.15 and $28.99 per minute. Other delay cost factors were for example maintenance and aircraft ownership (Airlines for America, 2022)

Conceptually, flight delays impact the financials of businesses in three ways. Firstly, delays affect the time customers spend at an airport. When customers have a longer waiting time, more customer support, maintenance, cleaning, and security are needed. Secondly, some customers may decide to fly from another airport or not fly at all due to extreme weather-related delays. Additionally, passengers experience the monetary effects of delays. They may face high prices of accommodation, transport, and re-bookings when they are required to change their travel plans on short notice. Flight delays can also lead to a loss in productivity, opportunity cost of leisure as well as missed business activities for passengers (Anupkumar, 2023). Finally, airlines face high financial impact as they need to spend more money on crew members, fuel, airplanes, and maintenance. Furthermore, airlines risk losing customers to the competition by not satisfying their service needs (Peterson, Kevin, Barczi, & Graham, 2013).

Important to mention is the variation in delay costs per airline. One reason for this is that low-cost carriers keep their operating expenses low, e.g., they pay their employees lower wages and have fewer service personnel. Therefore, the delay cost per minute is proportionally lower for budget airlines than for premium ones (Budd & Ison, 2020). Another reason for differences in delay costs by carrier is airlines' choices of how to handle a delay. For example, different costs will occur depending on how the flight management system will steer the aircraft: It can either fly faster to make up for the delay or fly at the normal speed to save fuel (Cook, Tanner, Williams, & Meise, 2009). An airline can also choose whether to wait for passengers or crew members that are late due to weather conditions (Cook, Tanner, & Lawes, 2012). All these options lead to differences in delay times between operating carriers.

2.3.2 Climate Risk Management for Airline Businesses

A key aspect of climate risk management in the context of this research is to handle the costs resulting from weather-related flight delays. Based on the profit formulas (1) and (2), there are several ways for airlines to handle these costs.

$$Profit = Revenue - Cost \tag{1}$$

$$Profit = Selling\ Price \times Quantity\ Sold - Cost \tag{2}$$

The first option is cost reduction. Potentially, this could be achieved by reducing weather delays. Since weather cannot be controlled, one approach to minimize costs associated with weather induced delays could be to schedule more flights at a time of the day that shows fewer weather-related delays or less weather-related delay cost. However, it has been hard to apply this so far because of tight flight schedules. Usually, planes tend to travel back and forth a specific route several times a day. Thereby, turnaround time is kept as short as possible as airlines only create revenue when planes are in the air and airports try to maximize gate usage. These short turnaround times also lead to more reactionary delays as there is no buffer in between flights to recover from delays initially caused by weather events. Hence, tight flight schedules and short turnaround times make it challenging to accommodate a higher number of flights at certain times of the day, e.g., when the forecasted weather is beneficial. Network managers already try to optimize flight plans by identifying capacity shortages, extreme weather events, etc., and adapt schedules according to their findings (Eurocontrol, 2018).

Another way in which airlines can accommodate the costs incurred by weather-related delays is lowering costs in other areas such as personnel costs, maintenance costs, airport fees, etc. Due to high competition in the market and changing customer demands, constant cost benchmarking and improvement has been a common practice in the past and is essential to remain competitive in the future (Koopmans & Lieshout, 2016).

Airlines can also deal with costs by proportionally increasing sales or prices of flight tickets or complementary services. Unit sales of flight tickets have been continuously increasing in the past with a drop in 2020 due to COVID-19. In 2023 the sales numbers are still recovering from the effects of the pandemic. However, forecasts project constant growth for future ticket sales (International Air Transport Association, 2022). In terms of ticket prices, the US Bureau of Transportation Statistics publishes past annual US domestic average itinerary fares. The statistics show an overall development of dropping airfares since 1995. Due to the pandemic, 2021 and 2022 represent the first years with rising prices compared to the previous years since 2014 (Bureau of Transportation Statistics, 2023).

Regarding services offered by airlines, there has been a shift from services being included in the ticket price to selling them separately. Low-cost carriers started to sell tickets without services such as checked luggage and catering. Instead, they introduced the option of additional charges for these services to the market. In the past few years, this trend has been adopted by some premium airlines as well. Airlines also generate revenue from traditional services, e.g., lounge access, as well as new services, e.g., personalized transportation services (Dennis, 2007).

Besides directly dealing with the delay costs, political cooperation seems to be of great benefit in the context of weather-related flight delay management. It could help with investments in making infrastructure more resistant to weather damage, building alternative travel routes, and implementing regulations to further prevent climate change (Zanni, Goulden, Ryley, & Dingwall, 2016). Cooperative organizations include for example the International Civil Aviation Organization, Eurocontrol, and the US FAA. Furthermore, political investment could help with technology updates such as improved weather forecasting systems and operational systems that help to better manage climate risk situations (Paraskevi Paraschi, 2023).

While airports have the best chance to minimize risk by obtaining resilient infrastructure and updated technology as mentioned before, airlines should place their focus on developing resilient flight schedules adaptable to uncertainties while keeping operational expenses as low as possible (Şimşek & Aktürk, 2022).

3 Data

3.1 Creation of the Dataset

The dataset used in this research combines three main data sources: flight delay data, weather data and cost data.

The first step was to collect flight delay data from the US FAA database. This database includes extensive data on all flights in the US for the past decades. To obtain the data needed for the specific analysis, a selection of variables was downloaded for the period from 2013-01-01 to 2023-08-31. Each row in the downloaded dataset represents one flight and includes the variables flight date (FL_DATE), operating carrier (OP_UNIQUE_CARRIER), departure airport (ORIGIN), original departure time (CRS_DEP_TIME), departure delay in minutes caused by weather (WEATHER_DELAY), departure delay in minutes caused by the National Air Traffic System which includes non-extreme weather delays (NAS_DELAY), whether the flight was canceled (CANCELLED), and the planned airtime (AIR_TIME) (Bureau of Transportation Statistics, 2023). The dataset was saved as a pandas data frame in a python-3 environment and contained 1,129,722 rows and 8 columns. After loading this dataset, it was carefully inspected and transformed: First, the data was filtered for flights that depart from JFK Airport only. The variable FL_DATE was turned into the datetime format to prepare it for modeling and further data wrangling. Next, the dataset was ordered by flight date from 2013-01-01 to 2023-08-31 and the index was reset. For modeling purposes, FL_DATE was later transformed into Day_of_Year. The inspection of null values followed, and all flights with NaN values for the variable AIR_TIME were removed as this value is needed as a predictor in the machine learning models. All NaN values of WEATHER_DELAY and NAS_DELAY were replaced by 0 as they mean that no delays were recorded for these variables. The dataset contained no other NaN values. Next, the binary variable Weather_Delayed was created with formula (3).

$$Weather\_Delayed(WEATHER\_DELAY, NAS\_DELAY) = \begin{cases} 1, & if\ WEATHER\_DELAY > 0\ and/or\ NAS\_DELAY > 0 \\ 0, & otherwise \end{cases} \tag{3}$$

A new categorical variable Day_Type was created indicating whether a flight was scheduled on a weekday or weekend. This provided additional information for the machine learning models in predicting whether a flight is delayed or not.

The next step was to collect weather data for the location of JFK Airport from two main data sources: historical ERA5 data and historical and future data from CMIP6. These weather variables were considered at daily scales, limited by CMIP6 availability.

Fifth and latest generation reanalysis ERA5 data was obtained from the ECMWF, available at the Copernicus Climate Data Store. A detailed description of ERA5 can be found in section 2.1.2. The variables "10m wind gust since previous post-processing", "Total precipitation rate since previous post-processing", "Maximum 2m temperature since previous post-processing", "Minimum 2m temperature since previous post-processing" were downloaded from the ERA5 dataset: For the wind gust and the total precipitation, the daily average, as well as the daily maximum value, were saved in a pandas data frame for each day in the period from 2013-01-01 until 2023-08-31. The daily maximum temperature and the daily minimum temperature two meters above the surface were collected in the same way. The unit of temperature was transformed from Kelvin to Degrees Celsius by formula (4).

$$Temperature_{Celcius} = Temperature_{Kelvin} - 273.15 \tag{4}$$

Each of these weather variables now existed in a separate data frame containing the Date column and the weather variable column. Next, each of the weather data frames was merged with the flight delay data frame, left on the FL_DATE variable and right on the Date variable, finally creating a new data frame including all variables of the flight delay dataset as well as the daily weather variables for each flight date. As the temperature values were only available until 2023-08-02, the new data frame was cut so it only contained values until the end of July 2023 (Copernicus Data Store, 2023).

CMIP6 NorESM2-MM data was obtained from Copernicus datastore, in particular historical CMIP6 data, SSP1-2.6 and SSP5-8.5 data. The historical data is available from 1850 to 2014 where real observations of climate exist. The historical model output is often used to evaluate the accuracy of a model by comparing it to real observations and can be used as a reference for future CMIP6 models. The future SSP scenarios are based on a change in the energy balance of the earth, measured in radiative forcing in watts per square meter. SSP1-2.6 represents a low radiative forcing of 2.6 W/m^2 and an increase in temperature of 1.7°C by the year 2100 (O’Neill, et al., 2016). The SSP1-2.6 scenario is in line with the Paris Agreement under the United Nations Framework Convention on Climate Change which restricts global warming to 2°C above pre-industrial levels by active climate protection (Kreienkamp et al., 2022). SSP5-8.5 on the other hand shows the most severe change in climate with a radiative forcing of 8.5 W/m^2 and a rise in temperature of 4.4°C or higher by the year 2100 (O’Neill, et al., 2016). The SSP5-8.5 represents intensive and increased use of fossil raw materials and energy, while limiting climate protection to a minimum (Kreienkamp et al., 2022). Therefore, using these two scenarios for the predictive models of this research will cover the entire range of expected climate change-related developments. The expected temperature change by the SSPs is displayed in Figure 1 (Frazier, Halpern, Vargas, & Lombard, 2022).

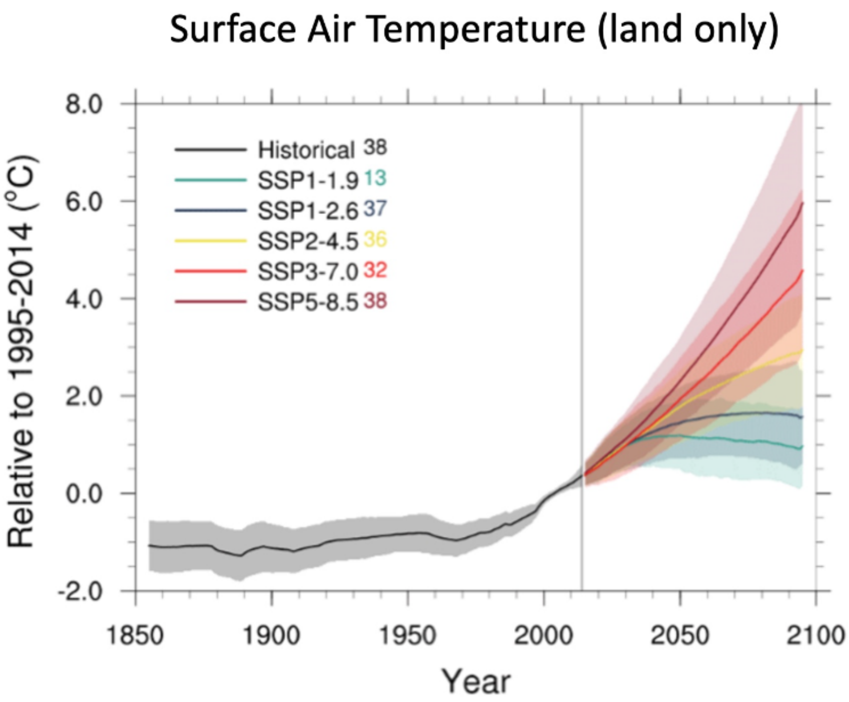

Figure 1: Temperature Change by Shared Socioeconomic Pathways (Frazier, Halpern, Vargas, & Lombard, 2022)

Based on these two scenarios, SSP1-2.6 and SSP5-8.5, predictions for yearly future flight delays were made to understand how different intensities of climate change will impact the airline industry. The first step was to identify variables in the CMIP6 dataset that compare to variables from the ERA5 dataset. The four variables daily average “near-surface wind speed”, “precipitation”, “daily maximum near-surface air temperature” and “daily minimum near-surface air temperature” were retrieved from the CMIP6 data by stating the coordinates of JFK Airport, a latitude of 40.6413 and a longitude of -73.7781. This was done to extract historical data from 2003 until 2014 due to limited availability and for future data from 2023 until 2030 for the two climate scenarios SSP1-2.6 and SSP5-8.5. As before, the temperature data had to be transformed from Kelvin to degrees Celsius by formula (4). All in all, for each future and past weather variable, a data frame containing the column date and the associated daily average value was created (Copernicus Data Store, 2021). The historical dataset was compared to the ERA5 observations to correct bias using Quantile Delta Mapping (QDM) (see section 4.2 for details).

Finally, the FAA estimated that each minute of flight delay led to an average cost of $101.18 for American airline businesses in 2022 (Airlines for America, 2022). Hence, the variable Estimated_Delay_Cost was created by applying formula (5).

$$Estimated\_Delay\_Cost = (WEATHER\_DELAY\ +\ NAS\_DELAY) \times \$101.18 \tag{5}$$

This variable will help to make business indications later on, e.g., by estimating how the weather-related delay cost will develop per flight and airline.

3.2 Variable Inspection

As mentioned in chapter 3.1, the binary output variable Weather_Delayed was created with formula (3). WEATHER_DELAY represents minutes of delay for which the reported cause was extreme weather. Additionally, it was decided to include NAS_DELAY into the outcome which reports minutes of delay due to the National Aviation System. It was stated in the variable description on the FAA website that this variable includes non-extreme weather delays (Bureau of Transportation Statistics, 2021). Therefore, the newly created output variable Weather_Delayed captured all flights that were delayed due to extreme as well as non-extreme weather. It was decided not to only look at the extreme weather delays because this would have increased the already existent imbalance between flights that were on time and that were delayed in the dataset. Figure 2 visualizes this imbalance of 762,128 on-time flights to 118,665 delayed flights.

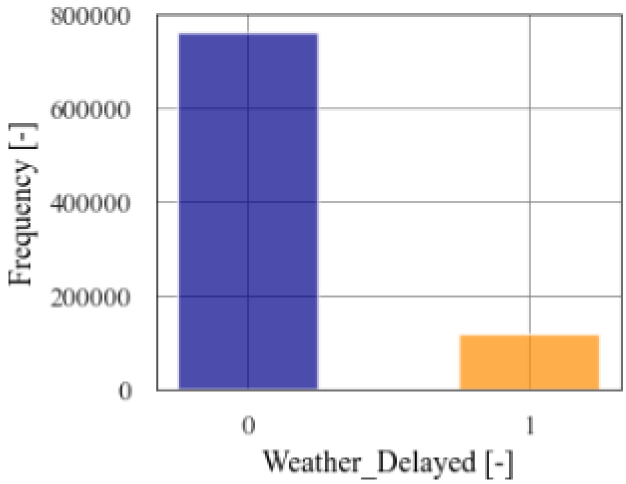

Figure 2: Number of On-Time Flights (blue) versus Number of Weather-Delayed Flights (orange)

Figure 3 shows the distribution of the four considered weather variables for on-time and delayed flights. The daily average wind gust variable ranges from 2.81 m s^{-1} to 20.57 m s^{-1}. The minimum temperature is in the -18.28 °C to 26.74 °C range, while the maximum temperature ranges between -10.98 °C to 34.38 °C. The average precipitation rate ranges from 0 kg m^{-2} s^{-1} to 0.00104 kg m^{-2} s^{-1}. Notice that Figure 3 hides outliers for enhanced visual clarity, hence, some maximum and minimum values are excluded from the below plots. The results showed slightly higher values for wind and temperature variables for flights that were delayed due to weather compared to on-time flights. Moreover, when observing the box plot in Figure 3, weather-delayed flights seem to have significantly higher precipitation values. Figure 3 visualizes how weather variables influence flight delays and therefore agrees with previous studies identified in section 2.2.1.

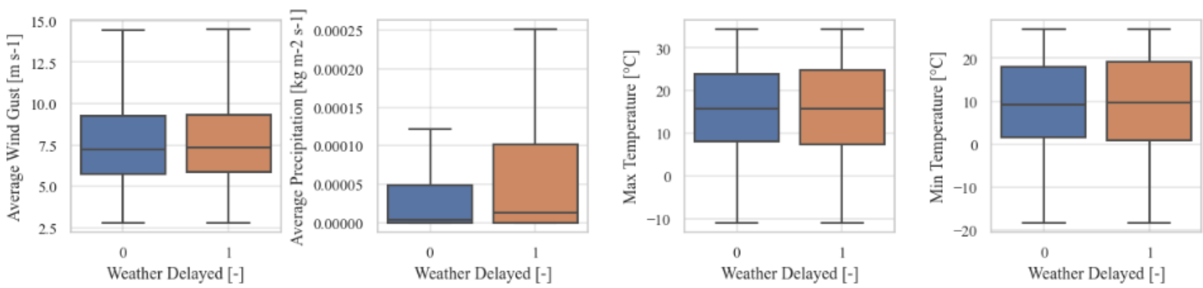

Figure 3: Distribution of Weather Variables for On-Time versus Weather-Delayed Flights

Besides the weather variables, time variables were included to predict whether a flight was delayed or not. One predictor is the Day_of_Year variable: As explained in chapter 2.2.2, the number of flight delays can be higher on specific days with increased air traffic such as holidays and popular vacation periods. The same is true for the Departure_Time and the

Day_Type: More popular flight times and flights during weekdays have proven to exhibit a higher probability of delay. Visual inspection of Figure 4 suggests that flights with longer airtimes have a higher chance of being delayed due to weather than the ones with shorter durations.

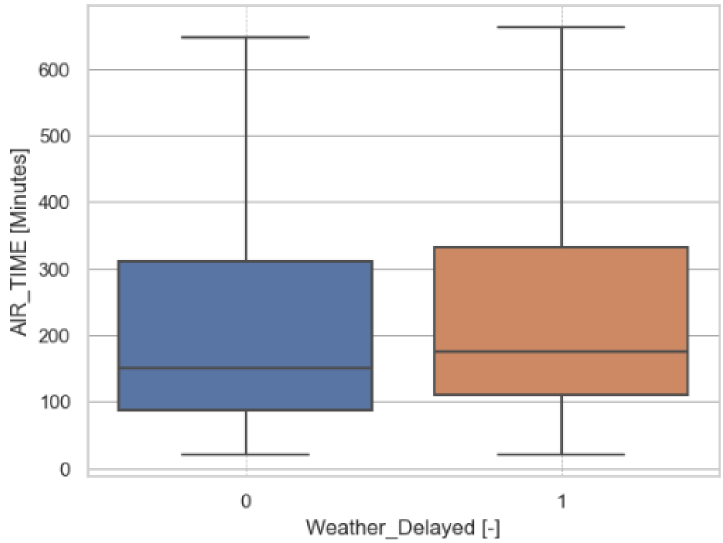

Figure 4: Distribution of Airtime for On-Time Flights versus Weather-Delayed Flights

Figure 5 shows that the number of flight delays varies with the operating airline. The airline HA has the lowest percentage of weather delays, while EV and MQ have the highest. Therefore, OP_UNIQUE_CARRIER will be considered a predictor variable for Weather_Delayed. The real airline names can be looked up via lookup tables on the website of the Bureau of Transportation Statistics (Bureau of Transportation Statistics, 2023).

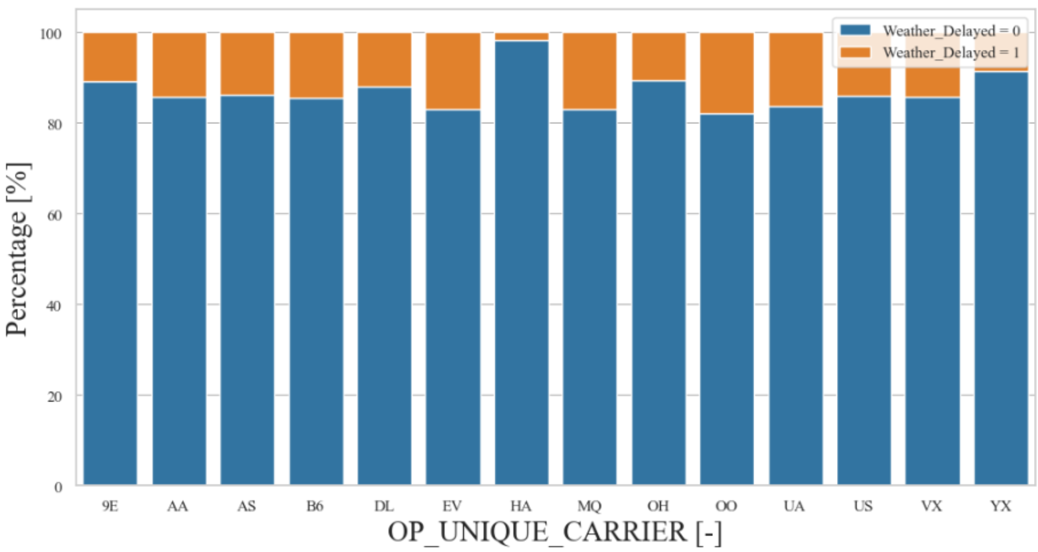

Figure 5: Percentage of Weather Delays per Operating Carrier

Before building a machine learning model, it is important to inspect the correlation between input variables. High correlation could introduce multicollinearity issues and lead to the violation of regression assumptions. Additionally, addressing the correlation between variables helps to reduce redundancy and bias and makes the model more interpretable. Therefore, a correlation matrix is shown in Figure 6. The most relevant result was the extremely high correlation of 0.97 between the maximum and the minimum daily temperature. Thus, the maximum daily temperature was removed before building a prediction model to decrease redundancy. Furthermore, there was a noticeable negative correlation between the average daily wind gust and the temperature, positive correlations between the daily average precipitation and the daily average wind gust, and between the daily temperature and the day of the year. Therefore, it was experimented with interaction terms between these variables in the modeling chapter with the goal of increasing prediction accuracy.

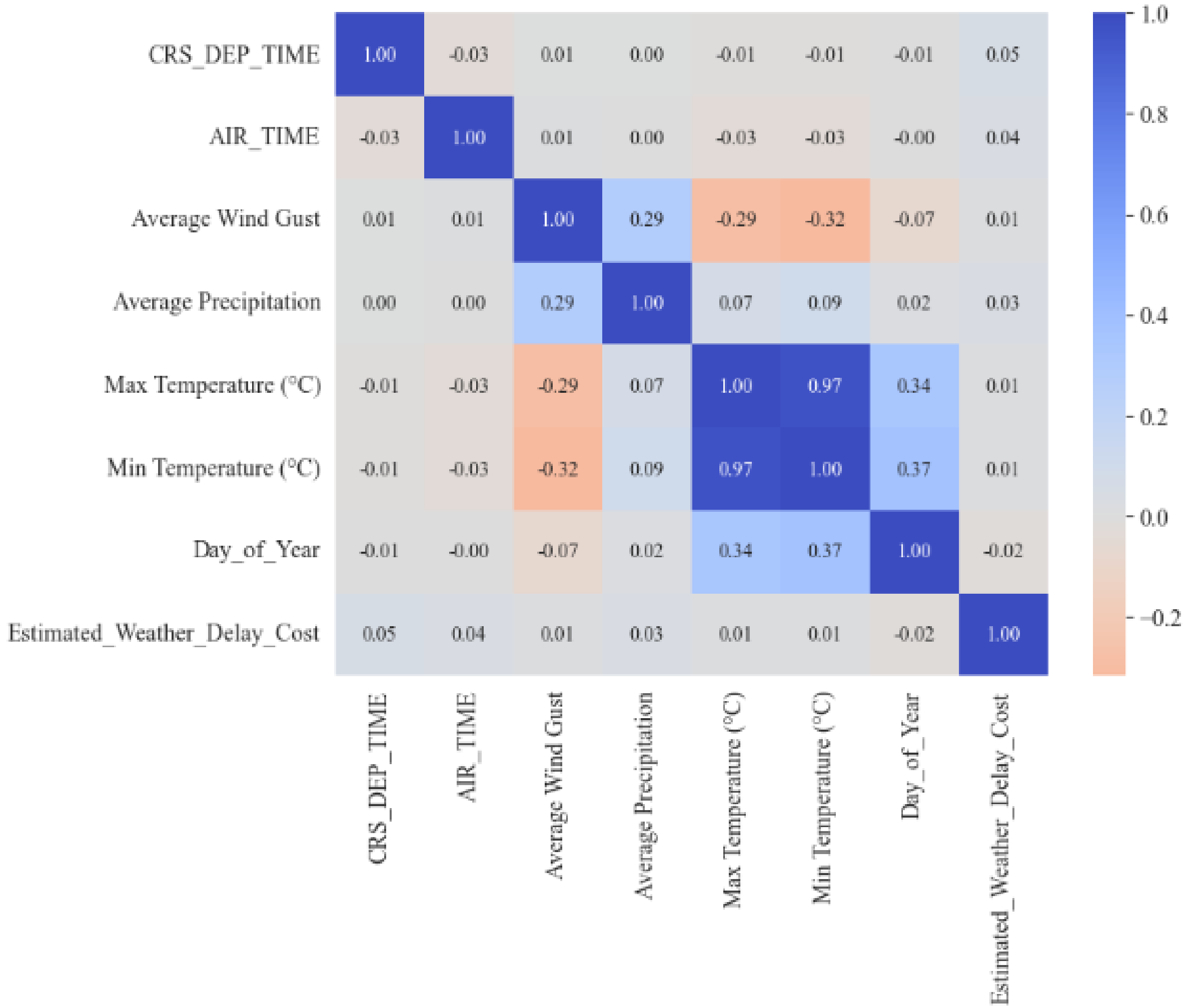

Figure 6: Correlation Matrix of Selected Variables

4 Methods

4.1 Modeling

In order to obtain the change in the number of flight delays caused by weather conditions from 2023 until 2030, first, a machine learning model was trained on the historical data to predict the binary outcome of whether a flight is delayed or on time. Three models of different complexity, all suitable for a binary classification task, were built and compared: a logistic regression model, a gradient-boosting model, and a random forest model. Furthermore, two regressor models were trained on the same data to predict the costs associated with each delay: a gradient-boosting regressor and a neural network.

4.1.1 Data Preprocessing

To make their performance comparable, the data was prepared in the same way for all classification models. First, the input and the output features of the model were defined: FL_Date, OP_UNIQUE_CARRIER, CRS_DEP_TIME, AIR_TIME, Day_Type, Average Wind Gust, Average Precipitation, Min Temperature (°C) were selected as predictors and Weather_Delayed as outcome based on the explanations in chapter 3.2. It was decided not to include Max Temperature (°C) due to the high correlation with Min Temperature (°C) discovered in chapter 3.2 to avoid bias of the model. Next, the data was separated into numerical and categorical columns. The numeric variables were then standardized using the StandardScaler function to guarantee a similar scale with a mean of 0 and a standard deviation of 1 for all numeric variables. The categorical variables were one-hot encoded using the OneHotEncoder function to turn them into binary vectors. This led to the creation of binary columns for each categorical variable and the dropping of the first category to abstain from multicollinearity. Finally, the standardized and the one-hot encoded features were merged back into one dataset prepared to be used in machine learning algorithms. To be able to evaluate the accuracy of predictions, a split into a training set of 80% and a testing set of 20% was defined for the preprocessed dataset and a random_state parameter was set to 42 to ensure the replicability of the code.

For predicting the delay cost, the gradient boosting regression model as well as the neural network prepared the data in a similar way like the classification. The same input features were used and the Weather_Delayed outcome was added to the input features. The output feature was defined to be the Estimated_Weather_Delay_Cost which was calculated as described in chapter 3.1. The variables were split into numerical and categorical, transformed by the StandardScaler and the OneHotEncoder functions, and merged back together. For the gradient boosting regressor, a k-fold object was used to split the data into train and test sets for each fold. The data was shuffled before each split and set to a random state of 42 for replicability. For the neural network, the data was split into 80% training and 20% testing data and set to the same random state of 42.

4.1.2 Logistic Regression Model

A logistic regression model predicts the probability of an outcome by applying a logistic function to transform the predicted values to be between 0 and 1. The downsides of a logistic regression are the linearity assumption that needs to hold and the inability to deal with complex variable relationships. Advantages, on the other hand, are computational efficiency and good performance in the case of close-to-linear relationships (Sperandei, 2014).

For the logistic regression, an undersampling strategy was applied to both the train and the test set due to the higher number of on-time flights than delayed flights that had been observed in 3.2. The sampling_strategy parameter was set to different values (e.g., 0.4, 0.5, 0.6, 0.7) to achieve the highest model accuracy when evaluating the model performance on the validation and test data. Two interaction terms were found to enhance performance and therefore added in the data processing part: one interaction between the encoded operating carrier variables and the departure time, and one between the operating carrier variables and the airtime.

Furthermore, it was experimented with different fixed hyperparameter values for the regularization (e.g., 10, 100, 1000) and the number of maximum iterations (e.g., 300, 500, 600, 1000). Additionally, stratified k-fold cross-validation with 3 folds was applied to ensure consistent model performance across several data subsets. The stratified version was chosen because of its ability to maintain the same distribution of delayed and on-time flights in each fold. This is important as the model was dealing with unbalanced data. A for-loop was used to iterate over each fold: During each iteration the dataset was split into the training and testing parts, the logistic regression model was trained on the training set and the model was used to make predictions on the test set. For each fold, the performance metrics accuracy, classification report, and confusion matrix were calculated and stored by comparing the true dataset values to the model predictions. The average values of all folds for each performance metric were calculated so they could be compared to the performance metrics of other models.

4.1.3 Gradient Boosting Classifier

The gradient boosting technique is also based on an ensemble of weak learners which are often decision trees. It continuously improves on past trees by correcting each tree by the residuals of the combined model. It adapts the weights of each tree and thereby reduces the loss function. Gradient boosting is known for high accuracy as it unites the insights of many weak learners, can deal with complex feature interactions, and looks at feature importance. The downsides are computational complexity, lower interpretability, and the tendency to overfit (Biau, Cadre, & Rouvière, 2019).

The gradient boosting classifier used the Synthetic Minority Over-Sampling Technique (SMOTE) to oversample the weather-delayed flight instances as this algorithm outperformed other sampling strategies. It was decided to refrain from interaction terms as no significant performance improvement could be observed. As in the logistic regression, a stratified k-fold cross-validation with 3 folds was implemented for the gradient-boosting classifier to ensure consistent performance. Here it was experimented with different combinations of values for the hyperparameters n_estimators (e.g., 100, 200, 300, 400, 600,1000) representing the number of decision trees, the learning_rate (e.g., 0.1, 0.2, 0.4) and the max_depth (e.g., 3, 5, 10) which controls the maximum depth of each decision tree. Lastly, a for-loop was applied in the same way as for the logistic regression and the average accuracy metrics were calculated.

4.1.4 Random Forest Model

Random forest applies ensemble learning and predicts classes based on several decision trees. It can also be used for regression problems by providing a mean prediction. Decision trees are based on a random selection of input variables and training data splits for each tree and the aggregation of the predictions of all trees. Random forest classifiers are beneficial in machine learning as they can reach high accuracy by the combination of several trees. They make predictions based on the importance of individual features and are comparatively resistant to overfitting. Some disadvantages are the computational complexity and biases towards the majority class, however, the last mentioned can be solved by various methods to address the class imbalance (Cutler, Cutler, & Stevens, 2012).

For the random forest model smote with a sampling strategy of 0.6 was applied to deal with the class imbalance after trying out several over- and undersampling strategies. As for the logistic regression, interaction terms between the encoded operating carrier variables and the departure time, as well as between the operating carrier variables and the airtime were found to be beneficial for model accuracy. Hyperparameter tuning for the random forest classifier included investigating different combinations of values for n_estimators (e.g., 100, 150, 200, 300) and the max_depth (e.g., 10, 20, 30, 40). In the same way, as for the other models, a stratified k-fold cross-validation with three folds was used, the same for-loop was applied, and average accuracy metrics were calculated.

4.1.5 Gradient Boosting Regressor

The Gradient Boosting Regressor has the same underlying algorithm as the Gradient Boosting Classifier in 4.1.3 but differs in its task and loss function. Here hyperparameter tuning was applied by experimenting with values for the n_estimators (e.g., 10, 100, 1000), the learning_rate (e.g., 0.04, 0.4), and the max_depth (e.g., 5, 10, 12). K-fold cross-validation was applied in a for loop. To measure model performance the means of the Mean Absolute Errors (MAE), the Mean Squared Errors (MSE), and R-squareds were calculated.

4.1.6 Neural Network

A neural network algorithm is inspired by the human brain and consists of connected nodes structured into layers. On the one hand, these models can learn complex patterns from extensive, high-dimensional data. On the other hand, neural networks require training on large amounts of data and are computationally intensive (Dashti Latif, et al., 2023).

To predict the delay cost based on the outcome of the classification model, different network architectures were tried out for the neural network: Models were run with zero, one or two hidden layers, with and without a dropout layer, with different numbers of neurons (e.g., 32, 64, 128), different activation functions (e.g., relu, sigmoid, tanh, huber) and different loss functions (e.g., huber, mean squared error). On top of that, it was experimented with the number of training epochs (e.g., 10, 20, 50), the batch size (32, 64) and manually setting a learning rate (e.g., 0.1, 0.01, 0.001). The Mean Absolute Error, the Mean Squared Error, and the R-squared were calculated to compare the model performance to the gradient boosting regressor.

4.2 Bias Correction

The weather variables output by climate models are affected by well-known biases compared to observations (Navarro-Racines, Tarapues, Thornton, Jarvis, & Ramirez-Villegas, 2020). Consequently, postprocessing of climate model outputs through the application of bias correction methods is generally a prerequisite step for most climate change impact studies (Mishra, Bhatia, & Tiwari, 2020).

Here the CMIP6 data was bias corrected to the ERA5 observational dataset by employing QDM, a widely used bias correction technique for climate applications, specifically designed to adjust future climate model projections so that the full statistical distribution becomes more consistent with the historical observations while preserving the trends across the different quantiles (Cannon, Sobie, & Murdock, 2015).

First, the 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9 quantiles were calculated for the ERA5 values of each climate variable as well as for the CMIP6 historical values of each variable over eleven years from 2003 to 2014. Next each quantile value of CMIP6 was subtracted from the according quantile value of ERA5 to calculate the delta between both datasets for each quantile. Next, the same quantiles were calculated for each variable for the CMIP SSP1-2.6 and the CMIP SSP5-8.5 future data over eleven years from 2019 to 2030. For each value in the future data, the quantile it belonged to was identified in a new column. Depending on which quantile each future value was in, the matching quantile delta of the past data was added to the future value. In this way, the future data was adapted to the ERA5 data.

4.3 Predicting on Future Data

To be able to make predictions on the corrected future values of the climate variables, these variables were combined with the flight data frame: Data from the year 2022 was extracted from the past flight delay data frame. For each year from 2023 until 2030, the corrected future climate data was matched on the Day_of_Year variable with the 2022 flight delay data. In this way each future year was considered to have the same flight schedule as the year 2022. Now for each future year, a complete data frame had been created, consisting of the same variables as the data the machine learning model was trained on. Therefore, predictions of flight delays for each future year from 2023 until 2030 for the two climate scenarios SSP1-2.6 and SSP5-8.5 could be made.

The best-performing machine learning model on the test data was used to make predictions on the future data frames of each year. The data was transformed in the same way as when training the model: The numerical variables were scaled to a similar range and the categorial columns were one-hot encoded to make them readable for the model. The model was then used to predict the class 0 or 1 for each row, indicating whether a flight will be delayed or not as well as the associated probabilities of each prediction. In this way, the number of delays could be forecasted for the coming years. The predictions were inspected and interpreted in chapter 5.

After forecasting whether each flight would be delayed or on time for the future data of each year, the classification results were used as input for the best performing model for cost prediction. The results data of the classification model was transformed in the same way as the training data for the gradient boosting regressor. Therefore, the gradient boosting regressor was then used to predict the delay cost for each row of the results data. Chapter 5 displays how this was continued by forecasting the total delay cost for each future year and how this can be interpreted.

5 Results

5.1 Modeling

5.1.1 Flight Delay Prediction Models

To predict future flight delays, the best-performing classification machine learning model had to be chosen. Therefore, the average performance of the logistic regression, the gradient boosting classifier, and the random forest classifier were compared. To give a holistic view of model performance, several metrics were computed using the test data:

- Accuracy which gives the overall correctness of the model and can be computed with formula (6)

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions} \tag{6}$$

- Precision which measures the accuracy of positive predictions (Weather_Delayed equals 1) by applying formula (7)

$$Precision = \frac{True\ Positives}{(True\ Positives\ +\ False\ Positives)} \tag{7}$$

- Recall which calculates the ability of a model to capture all the relevant instances of a positive class with formula (8)

$$Recall = \frac{True\ Positives}{(True\ Positives\ +\ False\ Negatives)} \tag{8}$$

- F1-Score which provides a balance between precision and recall and is computed as shown in formula (9)

$$F1 = \frac{2\ \times Precision\ \times\ Recall}{Precision +\ Recall} \tag{9}$$

The average accuracy of the logistic regression model returned a value of 0.64, while the gradient boosting model had an accuracy of 0.93 and the random forest of 0.88. The average precision displayed the value 0.60 for the logistic regression, 0.93 for the gradient boosting classifier, and 0.56 for the random forest. The mean of the recall values of the three folds returned 0.54 for the logistic regression, 0.93 for the gradient boosting, and 0.42 for the random forest model. The mean value of the F1-Score was 0.51 for the logistic regression, 0.93 for the gradient boosting, and 0.48 for the random forest model.

It can be observed that the logistic regression model performs poorly compared to the other two models. The variable relationships seem to be too complex to be captured by a simple logistic regression. Therefore, it is suggested to use a model that can capture more difficult connections in the data. Both random forest and gradient-boosting classifiers fulfill this requirement. However, the random forest model returned a low mean recall value and F1-Score. This indicates difficulties in predicting positive instances, so the majority of flight delays are not forecasted by the random forest model. In contrast, the gradient boosting model shows consistently high values across all performance metrics as well as the highest accuracy of all three models.

Therefore, the gradient boosting model was selected as the model to predict flight delays caused by weather. The best-performing gradient boosting classifier had 1000 n_estimators, a learning rate of 0.4, and a maximum depth of 5.

5.1.2 Flight Delay Cost Prediction Models

As a next step, the best model for predicting the cost of flight delays had to be identified. This was done by training the gradient boosting regressor and the neural network regressor on the output data of the best classification model and then comparing the results by computing the following metrics using the test data:

- The Mean Absolute Error (MAE) returns the mean of all absolute differences between the true cost of delay and the predicted cost of delay. The general expression can be found in formula (10). The lower the MAE, the closer the predictions are to the actual values. Hence, the lower the MAE, the better the model performs.

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| y_{i_{true}} - y_{i_{predicted}} \right| \tag{10}$$

- The Mean Squared Error (MSE) shows the mean magnitude of the differences between predicted and actual values and can be computed with formula (11). It is larger when the differences are larger. Therefore, a higher MSE indicates the existence of larger prediction errors.

$$MSE = \frac{\sum_{i=1}^{N}(y_{i_{true}} - y_{i_{predicted}})^2}{N} \tag{11}$$

- R-squared provides the proportion of variance in the outcome variable that can be explained by the input variables. Its calculation is displayed in formula (12). R-squared returns values between 0 and 1. The closer it is to 1, the greater is the proportion of the variance that can be explained by the model.

$$R^2 = 1 - \frac{\sum_{i=1}^{N} y_{i_{true}} - y_{i_{predicted}})^2}{\sum_{i=1}^{N}(y_{i_{true}} - y_{mean})^2} \tag{12}$$

The gradient boosting regressor returned a mean MAE of $324.61, while the neural network regressor had an MAE of $341.21. The average MSE of the gradient boosting regression model displayed 4,507,584.83, and MSE of the neural network was 4,914,674.32. While a R-squared of 0.28 could be obtained from the gradient boosting regressor, a value of 0.24 was given by the neural network.

Both models show similar values, however, the gradient-boosting regression model performed slightly better in predicting the weather delay cost per flight. The MAEs of both models indicate only an error of about 3 minutes of delay time (one minute of delay costs an average of $101.18 as found in section 2.3.1) and therefore suggest good model performance. The MSEs of both models seem to be high which means some prediction errors might be higher than others. The R-squared values both show moderate performance, nonetheless, there seem to be variables missing to improve prediction accuracy.

To summarize, the best-performing model was the gradient-boosting regressor with 10 n_estimators, a learning rate of 0.4, and a maximum depth of 12. Additionally, it was chosen to set the loss function to quantile and alpha to 0.5 which approximates a huber loss function and is therefore less prone to outliers.

5.2 Development of Flight Delays

This chapter will inspect the results that the best classification model, the gradient boosting classifier, generated. Figures 7 and 8 show the total number of flight delays caused by weather events per year. The grey line represents the historical values that the model was trained on. While the number of flight delays went down from 2013 with 17,160 delays until 2016 with 13,253 delays, from 2016 onwards delays rose continuously until 2019 to 16,223. Due to COVID-19 flight delays reduced in 2020 and 2021 to 3,210 and 7,190 delays. This was due to fewer flights in general, so flight delays occurred proportionally less than in previous years. However, in 2022 the delays rose above the pre-pandemic level to 19,406 and continued the prior development of a continuously growing number of weather-related delays.

The red dotted line in Figure 7 shows the mean of yearly weather-related flight delays at JFK Airport that was predicted using the CMIP SSP1-2.6 data. The values vary around the pre-pandemic level mean of 17,119 delays. The red area around the mean represents the spread at a standard deviation of 1. This means that about 68% of predicted values fall within 15,584 and 18,385 delays. The red dots show the predicted values for each year and the error margin of each prediction. However, it is not recommended to rely on the absolute forecasted values, as the model generated random spikes and drops in some years due to well know natural interannual weather variability (Lüthi, Cress, Davies, Frei, & Schär, 1996).

The blue dotted line in Figure 8 displays the average of yearly delays that were forecasted with the CMIP SSP5-8.5 data. The blue line is slightly lower than the red one in Figure 7, while the values still vary around the mean of 16,312 delays at a pre-pandemic level. The blue area around the mean represents the spread at a standard deviation of 1. This means that approximately 68% of predicted values fall within 15,529 and 17,095 delays. The blue dots again show the predicted values for each year and the error margin of each prediction. The forecasted CMIP SSP5-8.5 values have a wider spread as well as slightly larger error bars than the CMIP SSP1-2.6 prediction model.

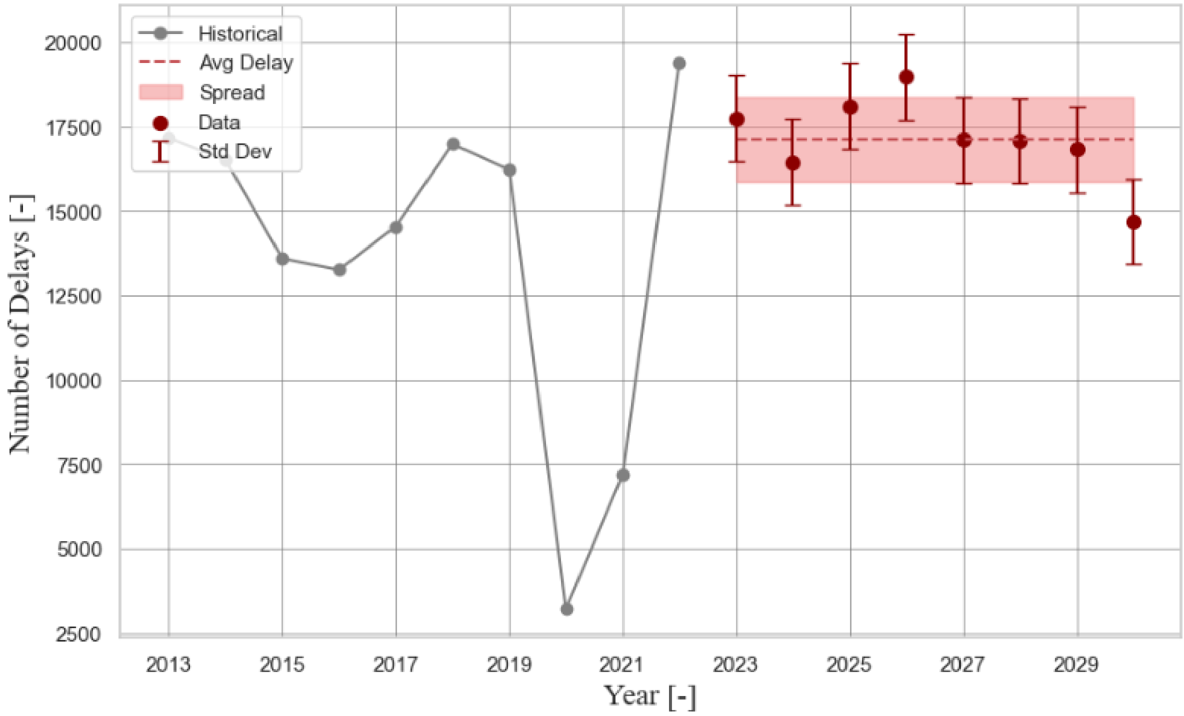

Figure 7: Weather Delays Timeseries CMIP SSP1-2.6 Data - Historical and Predictions

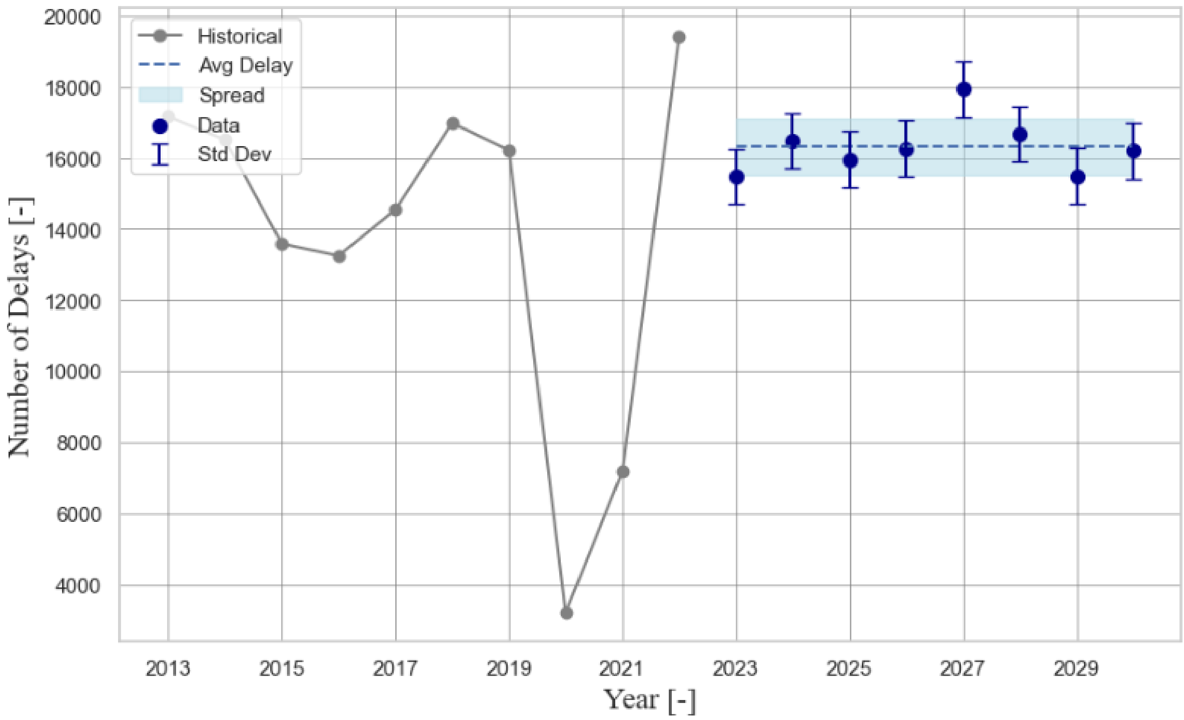

Figure 8: Weather Delays Timeseries CMIP SSP5-8.5 Data - Historical and Predictions

To interpret trends in the forecasted future values both the Kendall Tau and the Spearman Rank Correlation were calculated. They are measures that show the direction and the strength of the relationship between years and the number of weather-delayed flights. It was decided to calculate both measures and compare their outcomes to understand the robustness of the results. Kendall Tau is computed by counting how many variable pairs have the same order and how many have a different order and returns a value between -1 (perfect negative association), 0 (no association), and 1 (perfect positive association). Spearman Rank Correlation computes the correlation coefficient based on the difference between the ranks of the values of the two variables. Hereby, -1 represents a perfect negativ monotonic relationship, 0 means no monotonic relationship, and 1 means a perfect positive monotonic relationship. The p-values of the null hypothesis of no correlation/association were calculated for both measures, where a p-value of smaller than 0.05 indicates a statistically significant association/correlation.

For the forecasted number of yearly delays based on the CMIP SSP1-2.6 data, the Kendall Tau was -0.43 with a p-value of 0.18. This indicates a decreasing association between the variables year and number of delays. As the p-value is above 0.05, the association might not be statistically significant and the null hypothesis of no association between the variables cannot be rejected. The Spearman rank correlation for the CMIP SSP1-2.6 values was -0.5 with a p-value of 0.21. This is in line with the results of the Kendall Tau: There might be a negative correlation between the years and the number of weather-related delays. However, this is not significant at the alpha 0.05 level.

For the predicted CMIP SSP5-8.5 values, the Kendall Tau returned a value of 0.14 which indicates the number of delays increases over the years. Nonetheless, the p-value of 0.71 is above 0.05 and the null hypothesis of no association between the number of delays and years cannot be rejected. The Spearman Rank Correlation showed the value 0.19 which also means a positive monotonic relationship with a p-value of 0.65 which indicates non-significance.

All in all, the results for the future weather delays remain below the past values which means delay costs remain high but are not expected to grow markedly in the following eight years until 2030. While the results of Kendall Tau and Spearman Rank Correlation were not statistically significant, they show a slight decrease in weather-related flight delays for the CMIP SSP1-2.6 scenario. No observation in increasing climate change related delays is expected, due to the fact that this scenario only expects a small change in climate by 2100 and therefore only marginal change by 2030. On the other hand, the Kendall Tau and Spearman Rank Correlation showed a non-significant increase in the number of delays for the CMIP SSP5-8.5 scenario. This supports the assumption that a changing climate would lead to more weather-related delays. The insignificance of trends might be related to the small sample size of eight years or a signal-to-noise problem. While no significant increase in delays can be expected until 2030, these findings indicate that a more drastic increase in delays might happen further in the future.

5.3 Development of Flight Delay Costs

Interesting from a business perspective is the cost associated with the forecasted flight delays. Figures 9 and 10 show the total cost caused by weather-related flight delays per year. The patterns of the curves are similar to the flight delay curves above because the delay status is used as an input variable for the delay cost. The grey graph, representing historical values, shows a rise of cost from $45.52 million in 2015 to $72.99 million in 2019. COVID-19 made the cost drop in 2020 to $10.08 million, in 2021 to $27.6 million and in 2022 to $62.61 million.

In Figure 9, the red curve shows the CMIP SSP1-2.6 predictions. The costs vary around the average of $43.09 million and stay approximately constant. The red area around the mean represents the spread at a standard deviation of 1. This means that about 68% of predicted values fall within a cost between $39.12 and $47.07 million. The red dots show the predicted values for each year and the error margin of each prediction. However, it is not recommended to rely on the absolute forecasted cost values, as the model generated random spikes and drops.

In Figure 10, the blue curve plots the forecasted CMIP SSP5-8.5 values with an average of $41.99 million. On average, the SSP5-8.5 values are below the SSP1-2.6 values, however, it seems like the SSP5-8.5 values are higher in later years. As before, the blue area around the mean represents the spread at a standard deviation of 1. This means that approximately 68% of predicted values fall within a cost of $38.78 and $45.21 million. The blue dots show the predicted values for each year and the error margin of each prediction. As for the delay predictions, the forecasted CMIP SSP5-8.5 values have a wider spread as well as slightly larger error bars than the CMIP SSP1-2.6 values.

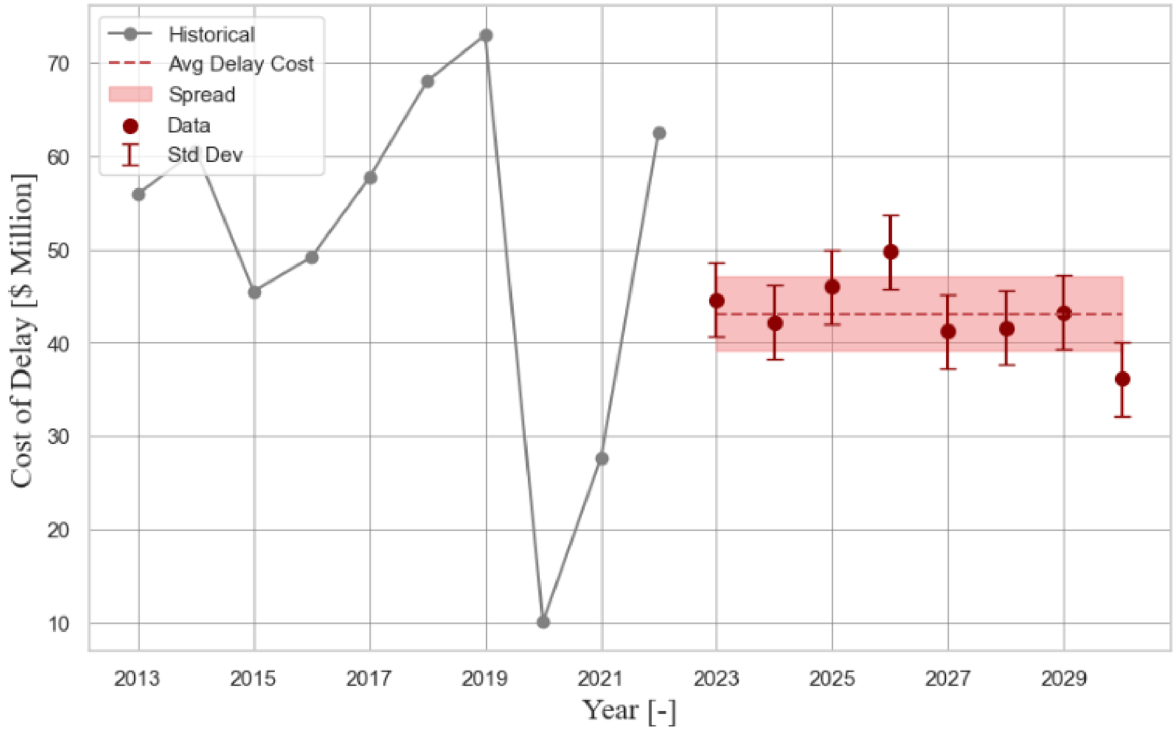

Figure 9: Weather Delay Costs over Timeseries - CMIP SSP1-2.6 Data

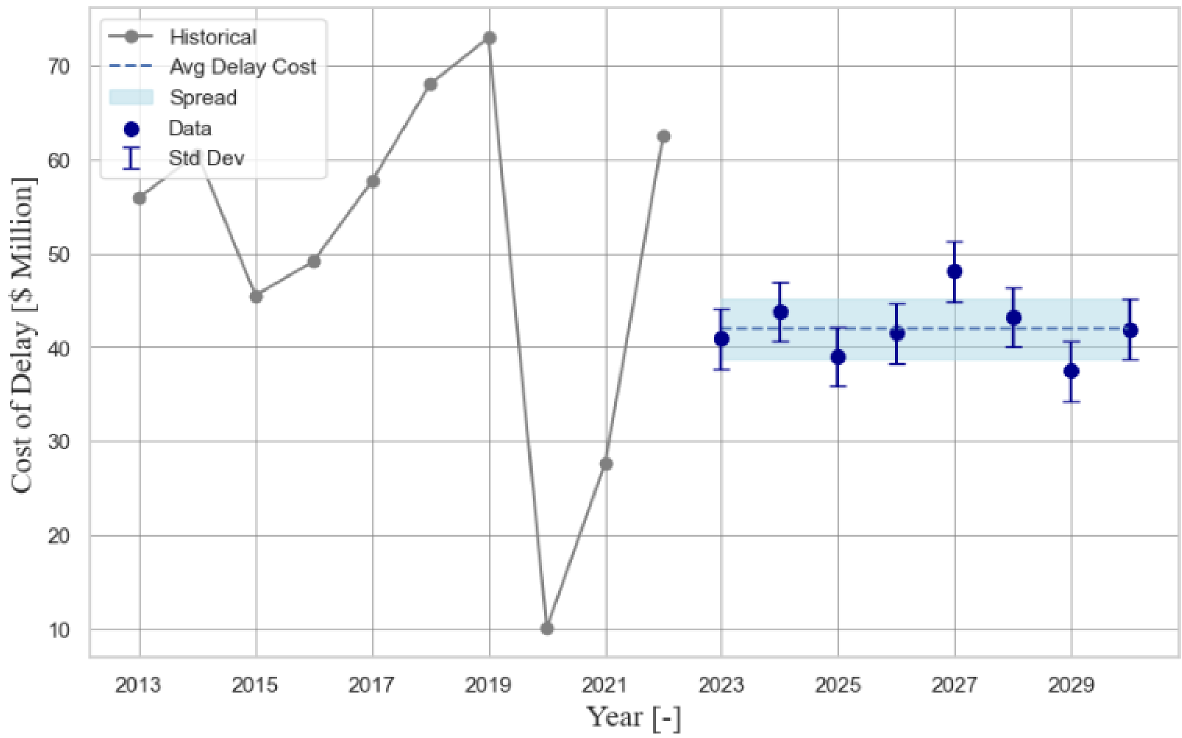

Figure 10: Weather Delay Costs over Timeseries - CMIP SSP5-8.5 Data

For the CMIP SSP1-2.6 data, the predicted yearly cost associated with weather-related delays had a Kendall Tau of -0.36 with a p-value of 0.28. This indicates a decreasing association between the features year and delay cost. As the p-value is above 0.05, the association does not indicate statistical significance and the null hypothesis of no association between the variables cannot be rejected. The Spearman rank correlation for the CMIP SSP1-2.6 values was -0.57 with a p-value of 0.14. This is in line with the results of the Kendall Tau: It implies a negative correlation between years and weather-related delay costs. However, this is not significant at the alpha equals 0.05 level.

For the predicted CMIP SSP5-8.5 values, the Kendall Tau had the value of 0.14 which shows an increase in costs over the years. Nevertheless, the p-value of 0.72 is above 0.05 and the null hypothesis of no association between the number of delays and the years cannot be rejected. The Spearman Rank Correlation displayed the value 0.24. This indicates a positive monotonic relationship with a p-value of 0.57, which is not significant.

To conclude, the results for the future weather-related delay costs remain below the past values which means delay costs remain high on average but are not expected to grow markedly in the next eight years until 2030. Even though the results of Kendall Tau and Spearman Rank Correlation were not statistically significant, they showed a slight decrease in weather-related flight delay cost for the CMIP SSP1-2.6 scenario. As for the number of delays, it makes sense to observe no increase in delay cost here. For the CMIP SSP5-8.5 scenario the Kendall Tau and the Spearman Rank Correlation show a non-significant increase in the delay cost. As before, the insignificance of trends might be related to the small sample size of eight years or to significant natural interannual variability that hides the trends. While no significant increase in delay cost can be expected until 2030, these findings indicate that a more drastic increase in cost might happen further in the future. Making predictions in the distant future will however deliver results with high uncertainty. Therefore, it is recommended to keep repeating the analysis conducted in this thesis every few years to continuously monitor the risk of an increasing number of delays.

6 Discussion

Based on the results discussed in chapter 5.3, this part will investigate the implications for airline businesses by using business analytics and general business principles. The first step of interest for the industry is identifying the most affected carriers by the cost of weather-related flight delays. As mentioned in chapter 2.3.1, there are several reasons for differences in delay costs between carriers.

Figures 11 and 12 show for the CMIP SSP1-2.6 and the CMIP SSP5-8.5 data which carriers were forecasted to have the highest average delay cost per flight. This was computed by summing the total cost by carrier and dividing it by the number of flights this carrier conducted. Based on formula (13), this was done for every year i and the average over all years was taken and displayed in the below figures.

$$Average\ Cost\ per\ Flight_{carrier} = \frac{\sum_i^N Total\ Cost_{carrier_i}}{\sum_i^N Number\ of\ Flights_{carrier_i}} \tag{13}$$

This ranking is supposed to inform airline businesses about their competitive position when it comes to delay costs at JFK Airport and can motivate the carriers to benchmark themselves to competitors with lower costs of delay. Both datasets generated the same ranking of players: UA faces the highest weather delay cost followed by AA or B6. In the middle are the carriers AS, and DL. 9E, YX, and HA face the lowest average delay cost per flight and should be used as a benchmark.

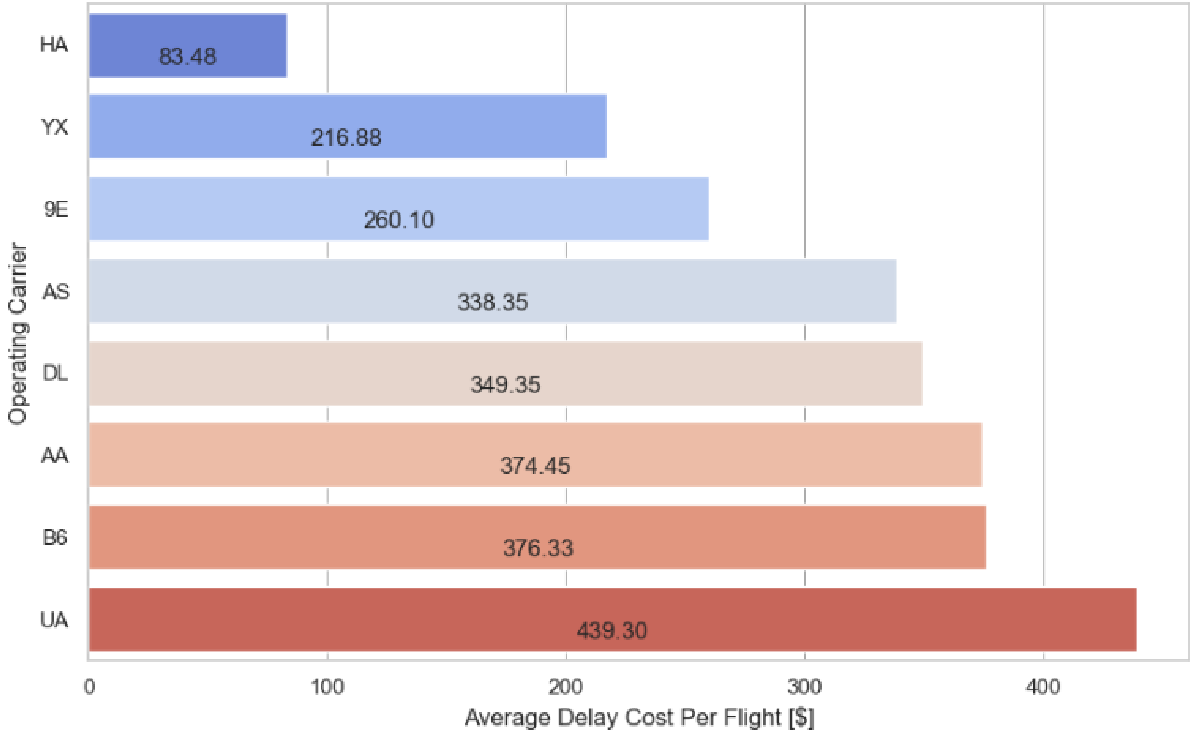

Figure 11: Average Delay Cost per Flight by Airline for SSP1-2.6 Data

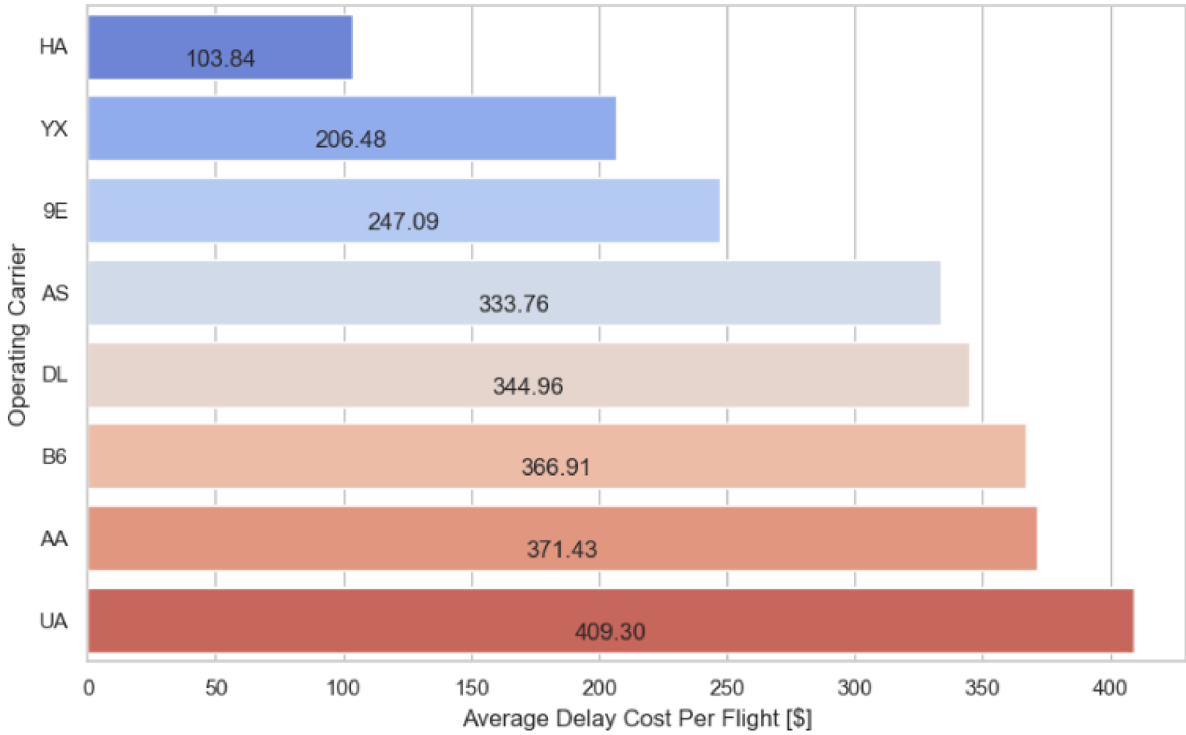

Figure 12: Average Delay Cost per Flight by Airline for SSP5-8.5 Data

A simplistic approach to dealing with rising costs due to weather delays can be used: The profit formula (1) and (2) as shown in chapter 2.3.2.

Based on this formula three approaches can be taken by the affected airlines to deal with the rising cost: cost reduction, increasing sales, increasing prices or a combination of these measures.

When it comes to lowering cost, one can either address the weather-delay cost itself or compensate for the higher weather-delay cost by decreasing other cost types. Lowering weather delay costs is tricky. One approach could be to schedule more flights at a time of the day that shows fewer weather-related delays. Figures 13 and 14 visualize that the average percentage of delayed flights of all scheduled flights varies by departure time, hence, it can be assumed that weather varies by time of the day as well. Both figures suggest rescheduling flights from between 1 pm to 10 pm to between 6 am and 1 pm. As found in chapter 2.3.1, however, it is difficult to implement improved scheduling as airlines try to maximize airtime and have tight flight schedules. What is more, flight plan optimization is already implemented by network managers.

Another approach might be to ensure robust and optimized infrastructure and machines as found in chapter 2.3.2. It is recommended to invest in the newest weather forecasting technology to adapt flight schedules before actual delays occur. Airlines are advised to stay as flexible as possible, for example by having more of their employees work on demand instead of on fixed shifts.

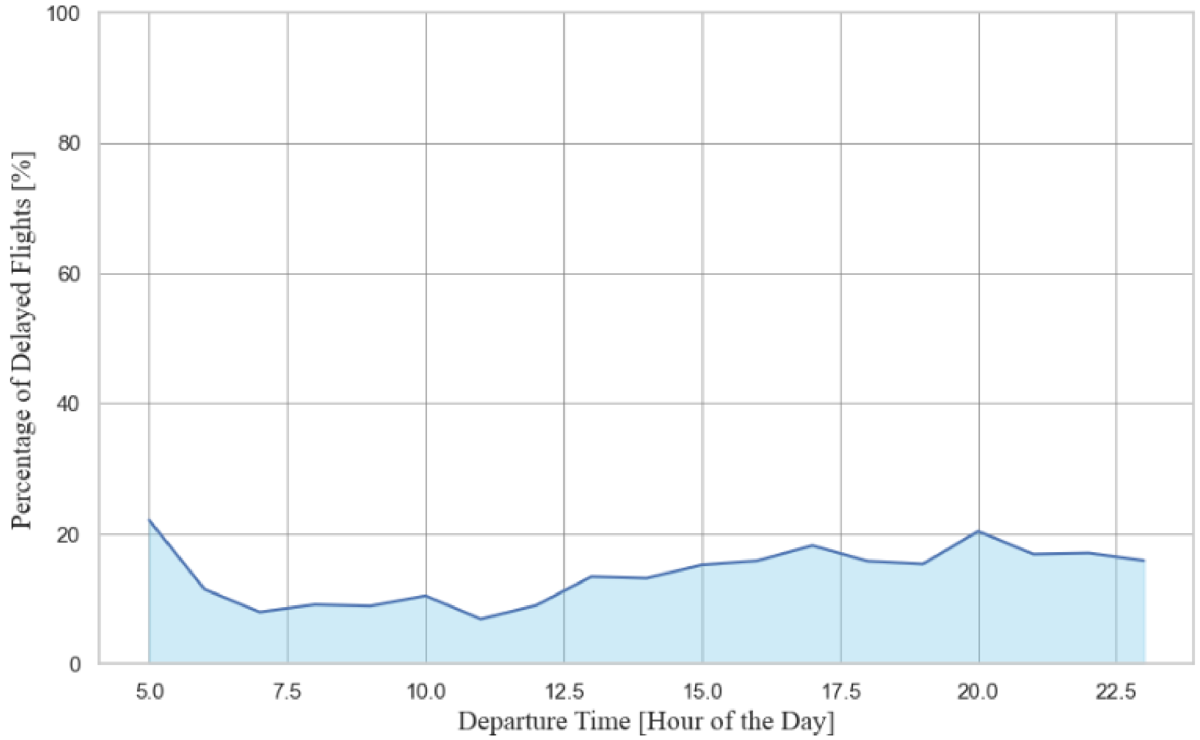

Figure 13: Mean Percentage of Delayed Flights over Departure Time – CMIP SSP1-2.6 Data

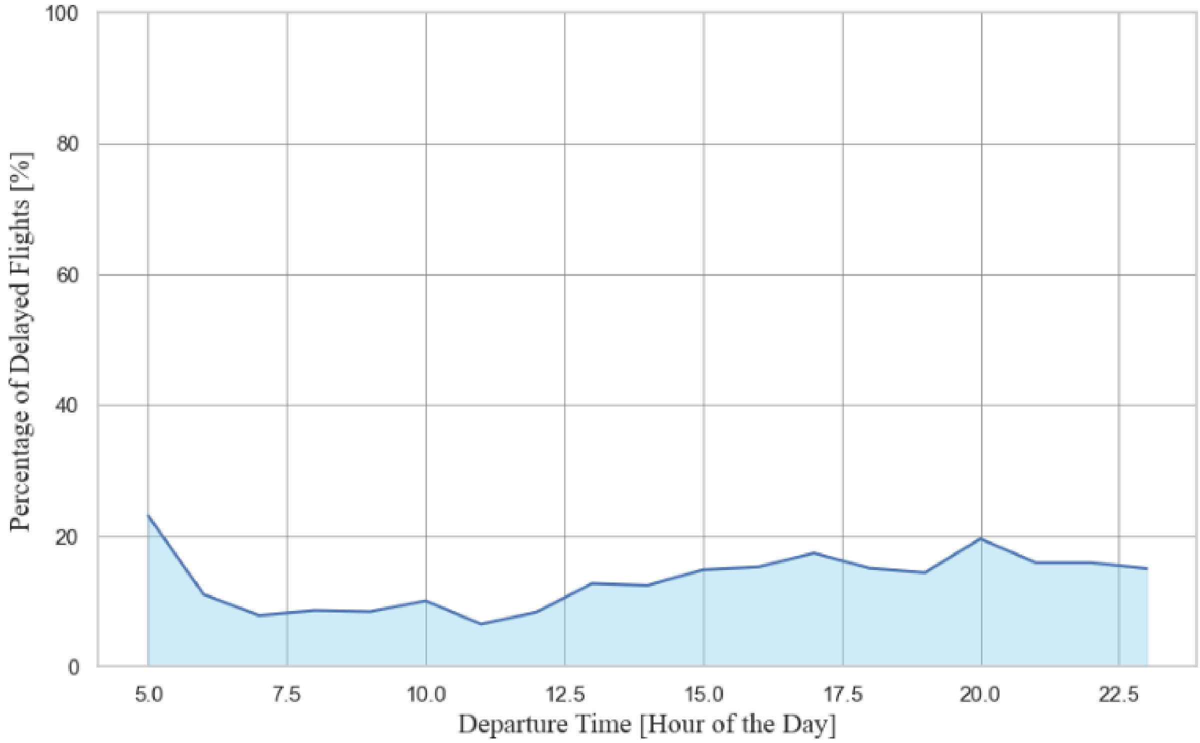

Figure 14: Mean Percentage of Delayed Flights over Departure Time – CMIP SSP5-8.5 Data

The third way to compensate for the higher delay costs is by lowering other cost types as much as possible. An airline can do this by benchmarking its costs, e.g., personnel costs, maintenance costs, airport fees, etc., to competitors and thereby identifying potential areas for cost reduction. Cost benchmarking has already been applied in the past and should be continued in the future to ensure the competitiveness of the business.

Based on the profit equation in formula (1) and (2), businesses can deal with higher costs by increasing sales. This can be done by either selling more flight tickets or selling complementary services. The cost for a flight ticket was on average $416.17 in 2022 for flights departing from JFK Airport (Bureau of Transportation Statistics, 2023). Assuming prices stay the same in the coming years this would mean that airlines need to sell 100,906 surplus flight tickets for flights departing from JFK Airport to fully compensate the average yearly cost for weather-related delays in the coming eight years until 2030 based on the CMIP SSP5-8.5 predictions. Based on the CMIP SSP1-2.6 data the change in sales would need to be 103,546 surplus tickets. The US Bureau of Transportation reported that 20,363,036 passengers departed from JFK Airport in 2022 on flights from domestic airlines (Bureau of Tranportation Statistics, 2023). This means ticket sales would have to be increased by 0.51% for the CMIP SSP1-2.6 data and by 0.50% for the CMIP SSP5-8.5 data. As found in chapter 2.3, the growing past demand for flight tickets shows a promising development for higher sales in the future. However, selling more flight tickets would mean scheduling excess flights which could in turn intensify the problem of flight delays. By introducing complementary services, e.g., airport lounge services, transportation from and to the airport, on-board streaming services, or entering new business areas such as hotel offerings, airlines could diversify and increase their sales and profits.

The last option of compensating additional weather-delay costs that can be drawn from the profit equation is an increase in prices. As before, two options are available here: increasing ticket prices and increasing the prices for complementary services. To accommodate the higher delay costs in the ticket price, the average ticket price would need to be increased by the average delay cost per flight of $304.78 for the SSP1-2.6 and $297.97.10 for the SSP5-8.5 data. Based on the average ticket price of $416.17 in 2022, this represents a price increase of 73.23% for the SSP1-2.6 and 71.60% for the SSP5-8.5 data. Such a high increase in prices is unrealistic, especially in consideration of the historical development of falling prices found in chapter 2.3.2. An increase in prices of complementary services, such as check-in luggage and meal offerings, could lead to customer dissatisfaction, as many of these services were offered for free before the COVID-19 pandemic.

This chapter gave an overview of business measures that could be taken to deal with the climate-related flight delay cost. Nonetheless, this represents just initial ideas and not detailed business measures. Each airline business is recommended to conduct a detailed and adapted analysis of what measures would be best to compensate for the delay costs in their specific case.

7 Limitations

One limitation identified in this thesis was the data availability. While the selection of weather variables was based on the literature review, their availability in the Copernicus datastore played a role, too. It was decided to obtain all variables from the Copernicus datastore as it represents a reliable and common source. By obtaining all weather variables from a single source, comparability of values is ensured. While broad availability can be observed for the ERA5 data, the CMIP6 data only includes very few variables. Additionally, the CMIP6 data does not contain data more granular than the day level. As showcased in chapter 2.2.2, the prediction of flight delays based on daily values is a valid approach. Another limitation might be that the CMIP data only includes average values but no extreme values for the variables wind gust and precipitation. While extreme values are often more insightful, it was shown that in this case extreme values would have made no difference: It was tried out to train the model on ERA5 daily extreme values (maximum daily precipitation and wind gust) as well as on daily averages but no significant difference in the model performance was observed.

The CMIP6 data was generated by only one model, NorESM2-MM, because processing multiple models at a higher resolution required a computing power that was not available for this study. This poses a constraint as a single model is only one plausible realization of the future, while multiple models would return a range of plausible outcomes for a given scenario.

Another major limitation was that the data used to train the machine learning model contained a historical period of only ten years due to computational capacity. This period was severely modulated by COVID-19. The number of flights in 2022 was still recovering from the pandemic, and it is expected that a full recovery will take a few more years. As the past data for this analysis stagnated at 2022 values, the projected future values are expected to represent an underestimation. While it was purposefully chosen to look at past data of the last ten years from 2013 until August 2023, historical CMIP data was only available until 2014. Hence, for the cause of QDM, a period from 2003 until 2014 was used instead of the same period as for training the model.

As explained in the chapter 3.1, when creating the outcome variable Weather_Delayed, all delays classified as NAS_Delay were included which means some of the delays included in the outcome variable might be due to other causes than weather. Nonetheless, this cannot be further improved due to a lack of documentation of the data source.

Worth mentioning is the exclusion of canceled flights as the dataset did not give any indication for the cause of cancelation and therefore the reason cannot be assumed to be weather.

Due to the scope of this work, it was decided to focus on specific input variables (time and weather variables), climate scenarios (SSP1-2.6, SSP5-8.5), delay types (departure), and one airport location (JFK Airport, New York City). Future research could build on the findings and look at other predictor variables (e.g., capacity), use different climate models, include more scenarios, include arrival and in-flight delays, or look at other or more airports.

Five different machine learning models were built and evaluated in order to select the most accurate ones for making predictions. Numerous other models could be built additionally or on top of these models.

Another limitation is the low availability of business data. Existing research is mostly limited to predicting flight delays but does not go further in offering business implications. Especially for airports limited qualitative recommendations can be found. Therefore, this thesis will offer new insights by combining the topic of flight delays caused by climate change with the science of business analytics. However, the quantitative focus will be on delay costs for airlines rather than for airports and individuals.

Lastly, the impact of the COVID-19 pandemic on the airline industry cannot be ignored. When inspecting the flight delay data, it can be observed that there were notably less flight delays in 2020 and 2021 due to a low number of flights in general. Therefore, when training the machine learning models it was tried to exclude these years from the data. However, they were added back as no difference in model performance could be observed. It can be inferred that, even though there were less flights in those years, the flights that took place had the same predictors for weather-related delays as in other years.

8 Conclusion

The first goal of this thesis was to investigate how the number of weather-related flight delays and the associated delay costs will develop in future years until 2030 compared to the past decade considering different future climate scenarios (CMIP SSP1-2.6 and CMIP SSP5-8.5). This was done by reviewing past literature on how climate change affects extreme weather, how climate modeling helps to forecast climate and weather patterns, how weather events can lead to flight delays, and what methods were used in the past to predict flight delays. These findings clarified what variables, data, and models would be best to generate the desired outcomes.

Flight delay data was combined with weather data and delay cost data to cover the aspects of hazard, exposure, and vulnerability for climate risk management of flight delays at JFK Airport under a changing climate in the next decade. The data was extracted, cleaned, merged, and prepared for modeling, and the considered variables were inspected. Several classification and regression machine-learning models were trained on the created coherent dataset. For the classification task of forecasting whether a flight was delayed due to weather or on time, a linear regression, a gradient boosting classifier, and a random forest model were created, and their performances were compared by accuracy, precision, recall, and F1-score. For the regression task of forecasting the delay cost for each weather-delayed flight, a gradient-boosting regressor and a neural network were trained on the historic data. Their results were compared through MAE, MSE, and R-squared values. The gradient boosting classifier and the gradient boosting regressor were found to be the best-performing models for predicting flight delays and the associated costs, respectively, and were therefore used as the basis for forecasting.

For the CMIP SSP1-2.6 scenario, the predicted average number of yearly weather-related delays for the eight years from 2023 until 2030 was 17,540 delays and therefore slightly above the pre-pandemic level. The average predicted yearly cost for weather-related delays for the same period was $45.29 million, similar to the yearly delay cost in the past. The Kendall Tau and Spearman Rank correlation found a non-significant decreasing trend in the predicted number of delays and delay cost. For the CMIP SSP 5-8.5, the results were similar, with an average of 16,623 delays at an average cost of $43.45 million per year. CMIP SSP5-8.5 showed a non-significant increasing trend in the number of delays and the associated delay cost.

To conclude, the above results showed no significant trends in the period from 2023 until 2030 compared to the last decade for JFK Airport. One reason for non-significance could be the short analysis period of eight years limited by the computing power. Another cause could be the presence of significant natural interannual variability that hides the trends due to a signal-to-noise problem. Lastly, the weather delay conditions at JFK Airport could be in fact not changing. It is recommended to regularly repeat the analysis of this research to spot the risk of increasing costs as early as possible and to be able to take timely precautions.

The second purpose of this research was to combine the disciplines of climate science and business analytics by establishing a climate risk management framework for an airline company with a focus on flight delays. This was initiated with the research of past literature on the pitfalls flight delays cause for airline businesses, and airlines' delay risk mitigation strategies. These findings from the literature were combined with the prediction results and the profit formula which is often used in business cases.

The results identified differences in the average delay cost per flight between airlines. Hence, it is advised for airline businesses to consider their specific business case and market position when handling flight delay costs and benchmark their cost mitigation strategy with the best-performing players. Based on the profit formula, it was proposed that airlines deal with the already high weather-related flight delay cost by cost reduction measures, e.g., by flight schedule optimization, cost benchmarks, and investment in up-to-date forecasting technology, or by increasing the quantity or the prices of plane tickets and complementary services.

As mentioned above, it is recommended to repeat this analysis every few years. The research can be extended by more historical data, a longer future period, more climate models with higher spatial and temporal resolution, additional weather variables, climate scenarios, delay types, and machine learning models. Future research could improve the cost modeling framework by adding more detailed data that was not available for this thesis. Additionally, airlines could model the global impact of flight delay predictions on their business and compare performance across locations and regions by including all airports they are operating at in the analysis.

References

Airlines for America. (2022). Retrieved 12 17, 2023 from U.S. Passenger Carrier Delay Costs: https://www.airlines.org/dataset/u-s-passenger-carrier-delay-costs/#:~:text=In%202022%2C%20the%20average%20cost,percent%20to%20%2442.15%20per%20minute.

al., K. e. (2022). *Empfehlungen für die Charakterisierung ausgewählter Klimaszenarien.*

Anupkumar, A. (2023). Investigating the Costs and Economic Impact of Flight Delays in the Aviation Industry and the Potential Strategies for Reduction.

Biau, G., Cadre, B., & Rouvière, L. (2019). Accelerated gradient boosting. *Machine Learning, 108*, 971–992.

Borsky, S., & Unterberger, C. (2019). Bad weather and flight delays: The impact of sudden and slow onset weather events. *Economics of Transportation, 18*, 10-26.

Budd, L., & Ison, S. (2020). *Air Transport Management. An International Perspective* (2 ed.). London and New York: Routledge.

Bureau of Tranportation Statistics. (2023, 08). Retrieved 12 12, 2023 from T-100 Market data: https://www.transtats.bts.gov/Data_Elements.aspx?Data=4

Bureau of Transportation Statistics. (2021, 03 16). Retrieved 12 12, 2023 from Airline On-Time Performance and Causes of Flight Delays: https://www.bts.dot.gov/explore-topics-and-geography/topics/airline-time-performance-and-causes-flight-delays

Bureau of Transportation Statistics. (2023). Retrieved 12 12, 2023 from Annual U.S. Domestic Average Itinerary Fare in Current and Constant Dollars: https://www.bts.gov/content/annual-us-domestic-average-itinerary-fare-current-and-constant-dollars

Bureau of Transportation Statistics. (2023, 08). Retrieved 12 12, 2023 from Average Domestic Airline Itinerary Fares: https://www.transtats.bts.gov/averagefare/

Bureau of Transportation Statistics. (2023, 09). Retrieved 12 17, 2023 from Reporting Carrier On-Time Performance: https://www.transtats.bts.gov/DL_SelectFields.aspx?gnoyr_VQ=FGJ&QO_fu146_anzr=b0-gvzr

Cannon, A., Sobie, S., & Murdock, T. (2015). Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? *Journal of Climate, 28*, 6938–6959.

Carvalho, L., Sternberg, A., Maia Gonçalves, L., Cruz, A. B., Soares, J. A., Brandão, D., . . . Carvalho, D. (2021). On the relevance of data science for flight delay research: a systematic review. *Transport Reviews, 41*(4), 499-528.

Christodoulou, A., & Demirel , H. (2018). Impacts of climate change on transport A focus on airports, seaports and inland waterways. A focus on airports, seaports and inland waterways. *Joint Research Centre Science for Policy report*.

Cook, A., & Tanner, G. (2015). European airline delay cost reference values. Updated and extended values. *University of Westminster*.

Cook, A., Tanner, G., & Lawes, A. (2012). The Hidden Cost of Airline Unpunctuality. *Journal of Transport Economics and Policy, 46*(2), 157–173.

Cook, A., Tanner, G., Williams, V., & Meise, G. (2009). Dynamic cost indexing – Managing airline delay costs. *Journal of Air Transport Management, 15*(1), 26-35.

Copernicus Data Store. (2021, 03 23). Retrieved 12 17, 2023 from CMIP6 climate projections: https://cds.climate.copernicus.eu/cdsapp#!/dataset/projections-cmip6?tab=overview

Copernicus Data Store. (2023, 12). Retrieved 12 17, 2023 from ERA5 hourly data on single levels from 1940 to present: https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview

Cutler, A., Cutler, R. D., & Stevens, J. R. (2012). Random Forests. In C. Zhang, & Y. Ma, *Ensemble Machine Learning* (pp. 157-175). New York: Springer.

Dashti Latif, S., Binti Hazrin, N., Koo, C., Lin Ng, J., Chaplot, B., Huang, Y. F., . . . Ahmed, A. N. (2023). Assessing rainfall prediction models: Exploring the advantages of machine learning and remote sensing approaches. *Alexandria Engineering Journal, 82*, 16-25.

Dennis, N. (2007). End of the free lunch? The responses of traditional European airlines to the low-cost carrier threat. *Journal of Air Transport Management, 13*(5), 311-321.

Dhadake, S., Shaikh, A., Sabale, S., Shinde, B., & Kadam, T. (2023). Flight Delay Prediction by Machine Learning. *International Journal of Advanced Research in Science, Communication and Technology, 3*(2).

Edwards, P. N. (2011). History of climate modeling. *Wiley interdisciplinary reviews: Climate Change, 2*, 128 - 139.

Eurocontrol. (2018, 08 03). Retrieved 12 12, 2023 from Delays – three questions and many answers: https://www.eurocontrol.int/news/delays-three-questions-and-many-answers

Fakhruddin, B., Boylan, K., Wild, A., & Robertson, R. (2020). Chapter 12 - Assessing vulnerability and risk of climate change. In J. Sillmann, S. Sippel, & S. Russo, *Climate Extremes and Their Implications for Impact and Risk* (pp. 217-241).

Federal Aviation Administration. (2020, 07 08). Retrieved 12 18, 2023 from Cost of Delay Estimates 2019: https://www.faa.gov/sites/faa.gov/files/data_research/aviation_data_statistics/cost_delay_estimates.pdf

Frazier, M., Halpern, B., Vargas, A., & Lombard, M. (2022, 01 28). *Ocean Health Index*. Retrieved 12 27, 2023 from CMIP6: PREDICTING FUTURE CLIMATE CHANGE: https://oceanhealthindex.org/news/cmip_1_what_is_this/

Frazzetto, D., Nielsen, T., & Šikšnys, L. (2019). Prescriptive analytics: a survey of emerging trends and technologies. *VLDB Journal, 28*, 575–595.

Frederikse, T., Landerer, F., Caron, L., Adhikari, S., Parkes, D., Humphrey, V., . . . Wu, Y.-H. (2020). The causes of sea-level rise since 1900. *Nature, 584*, 393–397. doi:https://doi.org/10.1038/s41586-020-2591-3

Gössling, S., Neger, C., Steiger, R., & Bell, R. (2023). Weather, climate change, and transport: a review. *Natural Hazards, 118*, 1341–1360.

Gratton, G., Williams, P., Padhra, A., & Rapsomanikis, S. (n.d.). Reviewing the impacts of climate change on air transport operations.

Gumus, B., Oruc, S., Yucel, I., & Yilmaz, M. T. (2023). Impacts of Climate Change on Extreme Climate Indices in Türkiye Driven by High-Resolution Downscaled CMIP6 Climate Models. *Sustainability, 15*(9).

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., . . . Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological society*.

Hertzberg, M., Siddons, A., & Schreuder, H. (2017). Role of greenhouse gases in climate change. *Energy & Environment, 28.4* , 530-539.

International Air Transport Association. (2022). *Air Passenger Market Analysis.* International Air Transport Association.

Johnson, R. (2019). The Costs of Flight Delays: Who Bears Them and How to Improve Performance. *Journal of Air Transport Management, 81*(101712).

Kim, Y., Choi, S., Briceno, S., & Mavris, D. (2016). A deep learning approach to flight delay prediction. *IEEE*, 1-6.

Koetse, M. J., & Rietveld, P. (2009). The impact of climate change and weather on transport: An overview of empirical findings. *Transportation Research, Part D 14*, 205–221.

Koopmans, C., & Lieshout, R. (2016). Airline cost changes: To what extent are they passed through to the. *Journal of Air Transport Management, 53*, 1-11.

Kumar, V., & Sherry, L. (2009). Airport Throughput Capacity Limits for Demand Management Planning. *Integrated Communications, Navigation and Surveillance Conference.*

L. Carvalho, A. S. (2020). On the relevance of data science for flight delay research: a systematic review. *Transport Reviews*.

Lui , G., Hon, K., & Liem, R. (2022, October). Weather impact quantification on airport arrival on-time performance through a Bayesian statistics modeling approach. *Transportation Research Part C: Emerging Technologies, 143*.

Lüthi, D., Cress, A., Davies, H., Frei, C., & Schär, C. (1996). Interannual variability and regional climate simulations. *Theoretical and Applied Climatology, 53*, 185-209.

Mishra, N., & Silakari, S. (2012). Predictive analytics: A survey, trends, applications, oppurtunities & challenges. *International Journal of Computer Science and Information Technologies, 3*(3), 4434-4438.

Mishra, V., Bhatia, U., & Tiwari, A. (2020). Bias-corrected climate projections for South Asia from Coupled Model Intercomparison Project-6. *Sci Data, 7*(1), 338.

Mokhtarimousavi, S., & Mehrabi, A. (2023). Flight delay causality: Machine learning technique in. *International Journal of Transportation, 12*, 230-244.

Navarro-Racines, C., Tarapues, J., Thornton, P., Jarvis, A., & Ramirez-Villegas, J. (2020). High-resolution and bias-corrected CMIP5 projections for climate change impact assessments. *Sci Data, 7*(1).

Neelin, J. D. (2011). *Climate Change and Climate Modeling.* Cambridge university press.

O'Neill, B., Tebaldi, C., van Vuuren, D., Eyring, V., Friedlingstein, P., Hurtt, G., . . . Sanderson, B. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development, 9*, 3461–3482.

Paraskevi Paraschi, E. (2023). Aviation and Climate Change: Challenges and the Way Forward. *Journal of Airline Operations and Aviation Management, 2*(1).

Peterson, E. B., Kevin, N., Barczi, N., & Graham, T. (2013). The Economic Cost of Airline Flight Delay. *Journal of Transport Economics and Policy, 47*, 107-121.

Pörtner, H.-Ö., Roberts, D., Tignor, M., Poloczanska, E., Mintenbeck, K., Alegría, A., . . . Langsdorf, S. (2022). Climate Change 2022: Impacts, Adaption and Vulnerability. Contribution of Working Group II to the SIxth Assessment Report of the Intergovernmental Panel on Climate Change. *Cambridge University Press*, 3056.

Rowley, C. R. (1998). Financing Airport Capital Development: The Aviation Industry's Greatest Challenge. *Journal of Air Law and Commerce, 63*(3).

Roy, D., Srivastava, R., Jat, M., & Karaca, M. (2022). A Complete Overview of Analytics Techniques: Descriptive, Predictive, and Prescriptive. In P. Jeyanthi, T. Choudhury, D. Hack-Polay, T. Singh, & S. Abujar, *Decision Intelligence Analytics and the Implementation of Strategic Business Management* (pp. 15–30).

Şimşek, D., & Aktürk, M. S. (2022). Resilient airline scheduling to minimize delay risks. *Transportation Research Part C: Emerging Technologies, 141*.

Sivakumar, S. (2022). A novel Integrated Risk Management Method for Airport operations. *Journal of Air Transport Management, 105*.

Škultétya, F., Jarošová, M., & Rostáš, J. (2021). Dangerous weather phenomena and their effect on en-route flight delays in Europe. *Transportation Research Procedia , 59*, 174–182.

Slama, B. R. (2016). Green house effect vs. infrared radiation emissions. *Journal of Climatology & Weather Forecasting, 4*(1).

Sovacool, B. K., Griffiths, S., Kim, J., & Bazilian, M. D. (2021). Climate change and industrial F-gases: A critical and systematic review of developments, sociotechnical systems and policy options for reducing synthetic greenhouse gas emissions. *Renewable and Sustainable Energy Reviews, 141*, 110759.

Sperandei, S. (2014). Understanding logistic regression analysis. *Biochemia medica, 24.1*, 12-18.

Tebaldi, C. (2020). Climate model projections from the scenario model intercomparison project (ScenarioMIP) of CMIP6. *Earth System Dynamics Discussions 2020*, 1-50.

Zanni, A. M., Goulden, M., Ryley, T., & Dingwall, R. (2016). Improving scenario methods in infrastructure planning: A case study of long distance travel and mobility in the UK under extreme weather uncertainty and a changing climate. *Technological Forecasting and Social Change, 115*.

Appendix 1 – Dissertation Code

Under the following link the complete code for the empirical part of this thesis can be found:

https://github.com/AnnaWimmer/Dissertation-Code