Minimizing Stress in Air-Travel

Elizabeth Hocking, Lauren Rock and Kailey Van Maren

April 2020

Contents

1 Introduction and Modelling Stress

The purpose of this paper is to investigate how to minimize the stress that an individual experiences during air travel from San Francisco to New York City. While there is no way to predict whether a traveller will definitively encounter a stressful situation, a model can be used to minimize the probability of experiencing stressful events. In the scholastic paper *Mathematical Modeling of Stress Management via Decision Control* by Matthew Shanahan, the researcher describes "stress related predictions are not necessarily of a discrete format, as they are in the case of Bernoulli-like win-loss occurrences. Rather than discrete outcomes to stimulus encounter, the probability of a stressing event more likely comes to the fore" (Shanahan, 2015). In the context of our paper, the stressful events are the binary variables assigned a 1 if they occur and a 0 otherwise. This does not interest us so much as the probability of each of these occurring. Is it worth risking a stressful situation if the chances of it actually occurring are slim? This is a classic case of Decision Control (DC). DC is a means of coping with stress by "positioning oneself in a multifaceted stressing situation so as to minimize the probability of an untoward event" (Shanahan, 2015). In our model we are trying to minimize the probability of a collection of events occurring, such as flight delays and cancellations. Through the creation of our model we utilized an interdisciplinary study of psychology, economics, statistics and optimization.

2 Economics of Decision Making

Through analyzing how economics impacts our paper we looked at consumer decision making when faced with an uncertainty of potential adverse events. In a paper titled *Are Consumers Strategic? Structural Estimation from the Air-Travel Industry*, the author examines how consumers make choices in this industry. We hope to further this research by adding in the complexity of uncertainty. When a consumer purchases a flight at a given cost, they make an assumption on what their welfare will be. Customers are not accounting for the potential reduction in consumer surplus if they encounter one of the stressful events our model covers. We hope that our model can assist customers in making more holistic decisions by providing them with information about how stress could potentially affect the outcome of their choice.

3 Psychology of Stress and Determining Weights

There is no way to objectively quantify stress, as the effects of stress varies from person to person across situations; the psychological aspect also needs to be researched to account for this variation. This has been one of the bigger struggles in the formulation of the model. Part of what intrigued us with this problem is the fact that so much of our lives is unquantifiable, and there has been little investigation into how we could try to scientifically assign values. We researched different decisions that people make to maximize their well-being or minimize their stress. We found that stressful situations are perceived differently as described by the different person-environment fit theory. The worse the person-environment fit, the more stressed a person is likely to be. In order to account for this in our model, we made the stress weightings in our Python files to be tunable by the user. The more concerned a person is about flying and the various things that could go wrong, the higher the weights they can assign to that factor. While we determined that calculating stress coefficients from primary research was outside the scope of the course, we did put some additional time into how we would prepare our survey had we had the opportunity to gather a significant amount of data.

4 How We Would Survey for Stress Coefficients

We decided to choose our coefficients for stress based on personal experience and general non-academic sources on the internet. It should be noted that our group did look into determining how these could be calculated statistically to represent the general population. To do this, a survey of a sufficient sample size, ideally more than 1000, should be randomly selected to present the list of the stressful factors that our model considers. Participants should then list all N stressors in order of most to least stressful $\{1, 2, ..., N\}$. This ordered-listing method would ensure that each stressor is included in the ranking and forces participants to evaluate each stressor on how much anxiety it would cause. Ordinal regression could then be used to calculate the coefficients. This method was chosen over both a Bernoulli experiment where participants would give stressors either a 1 if they perceive it as stressful or a 0 otherwise or a scale ranking surveying method. The problem with the former is we assume that the majority of people would find most, if not all of the stressors in the model to be stress-inducing. This would result in many of the stressors having the same weighting. The latter would likely have a negative skew, with most people's responses for most of the stressors falling between 7 and 10. Both of these would not fulfill our initial intentions of determining which events are most stressful, and would hinder our ability to calculate coefficients that represent the entire population's true perception of the stressors. As such, we leave it to individuals to determine their own individual preferences.

5 Flying Stressors

The probability that a flight would be delayed, cancelled, or diverted varies upon which airline and airport would be used as well as its scheduled departure time. A flight is deemed "on-time" if it arrives within the 15 minute window of its originally scheduled arrival time. Given our data set, we leveraged our background in statistics in order to manipulate the data into a working form for our program. The probabilities for the airlines are also based upon marketing carrier vs. the operating carrier, the difference being marketing carriers also include its partner airlines. For example a traveller may purchase a ticket through Delta, but the flight is actually operated by SkyWest Airlines. The model would capture the impact that partner airlines have on Delta's overall performance allowing for a more accurate representation that consumers face.

The data which contains the probabilities that each stressor occurs was collected from the December 2019 issue of Air Travel Consumer Report from the Office of Aviation Enforcement and Proceedings. The Consumer Report gathered and compiled information from the Bureau of Transportation Statistics. The probabilities were generated based upon all the flights that happened during the month of October 2019. It is important to note the time of the year the data was collected from as the travel industry could be subjected to seasonality biases. For example in July, a potentially busier month, airlines and airports may become disproportionately busier causing more delays which in turn could affect the optimality of our solution. The Consumer Report is released monthly so the model could be updated going forward to stay current.

6 Optimization Research in the Air Travel Industry

Our research into this field consisted of gathering airline industry data from commercial websites, statistical databases and news outlets. One of the goals of our study was to perform some statistical analysis on the real-world problem of travel stress. The research we conducted in the field of optimization was focused around what is currently being done to optimize the airline industry. Currently, there is a large focus on what organizations can do to optimize costs, flight patterns and scheduling. One paper by the title of *Optimization Applications in the Airline Industry* by Gang Yu and Jian Yang at the University of Austin Texas provided a comprehensive list of industry practices including Network Designs, Yield Management, Flight Planning and Fleet Assignment, Crew Scheduling and Air Traffic Flow Control. Some sample models used for these problems can be found in the exhibit section. These models were too complex for the scope of this project but helped us develop some of the assumptions for our model. In terms of the optimization research available from the consumer standpoint, the majority of it pertained to ticket pricing. The paper *Optimizing Airline Ticket Purchasing* by William Groves and Maria Gini of the University of Minnesota aimed to save customers money when purchasing tickets by determining the timeframe in which tickets should be purchased. Another paper that discussed potential applications of optimization in the consumer travel industry was by a group of 9 researchers in the Netherlands by the title of *Where to Go on Your Next Trip? Optimizing Travel Destinations Based on User Preferences*. Our research helped validate a potential opportunity to answer a question that had not yet been discovered and helped shape our motivation for the project.

7 Motivation

The motivation for this model breaks down simply to the fact that all three of us have had our share of teeth-grinding travel experiences, as most adults likely have. On a personal level, we are motivated to find ways to reduce the chances of stressful events, such as a cancelled flight, occurring to us. However, on a deeper social level we have found in researching the topic further that there are few sources today that investigate consumer behaviour around flight purchases, and especially one that analyzes it from a stress perspective. In an academic paper examining strategic consumer behaviour in the airline industry, the author stated “There is very little research to date that provides direct and rigorous evidence of (strategic consumer) behavior, relying instead on anecdotal accounts. Empirical evidence of strategic consumer behavior would not only enrich our knowledge of consumer behavior, but also improve managerial decisions" (Li, 2014). This quote and the rest of the paper were used in helping determine how the solution to our problem could be applied to administrative settings. Furthermore, we created a model that would not only be important for airline companies to use to understand the behaviour around what motivates flight purchases, but also to help consumers become more informed in their travel decisions. This emphasizes the value of this model to society, as it has many applications for different bodies of users, such as airlines, businesses, and individual consumers, and there is not yet widely available research on it.

8 Application of the Model

This model is valuable as there are both many different potential users, but also each user may use the model to achieve different ends or gather different insights. The user we had in mind when creating this model was the individual consumer, as our own experiences as travelers had inspired this topic of research to begin with. This model can be used by consumers before purchasing a flight ticket by inputting their preferences, running the model, then receiving output detailing which flight they should take to their destination based on their preferences. Furthermore, they could run the maximized version of the model and interpret the results to eliminate flights that would cause them the most stress when traveling.

Another application allows airline companies to understand their own performance in minimizing stress, and also how consumers make decisions when purchasing plane tickets. As consumers use this model, airlines will be able to analyze input data for which stressor weights are highest to determine which stressors are most important to limit in order to improve the consumer's perception. On top of that, airlines could analyze the results of the model and determine how different stressor weights affect their own performance in minimizing the stress of their customers.

Finally, businesses could use this model to minimize the stress their employees face when traveling for business. In recent years, there has been more investigation into the impact that stress has on work importance, and the findings have been that stress does reduce the quality of work performance and job satisfaction (Ivancevich, 2003). Thus it is valuable to businesses to reduce stress for a planned business trip, which can be done through utilizing our model. Businesses can also seek to minimize the likelihood of events that would be a detriment to the business trip, depending on their needs. Many business trips may be time sensitive, and thus it is especially important to reduce the likelihood of a late flight. A business user would find different events more or less stressful compared to a for pleasure traveller so the model allows for the customization of the stressor weights to account for this reason. Our model is valuable to an array of users in the industry because there is very little of it in the world now and it is useful for travelers to have this information.

9 LP Formulation

9.1 Constants

Stress weights are a constant value for each of the different stressors in our model.

$$\begin{aligned} S_l &= \text{Stress factor of flight being late} \\ S_c &= \text{Stress factor of flight being cancelled} \\ S_d &= \text{Stress factor of flight being diverted} \end{aligned}$$

In our python model we have default stress values at the top of the file, where

$$\begin{aligned} S_c &= 10 \\ S_l &= 5 \\ S_d &= 2 \end{aligned}$$

These values are customizable in the file to better reflect the user's personal travel preferences. These values come from our data table.

$$\begin{aligned} OnTime_{i,j} &= \text{the percent likelihood that a flight from airport i on airline j will be on time.} \\ Timely_{i,r} &= \text{the percentage likelihood that a flight from airport i at time r is on time.} \\ Diverted_{i,j} &= \text{the percentage likelihood that a flight from airport i on airline j will get diverted.} \\ Cancelled_{i,j} &= \text{the percent likelihood of a flight from airport i on airline j will get cancelled.} \end{aligned}$$

9.2 Variables

We have the following binary variables in our model:
$Airport_i$ represents which of the 3 NYC airports the flight leaves from.

$$Airport_i = \begin{cases} 1 & \text{flight leaves from airport } i \\ 0 & \text{otherwise} \end{cases}$$

$$i \in \{EWR, JFK, LGA\}$$

$Airline_j$ represents the airline which the flight is operated under.

$$Airline_j = \begin{cases} 1 & \text{if flight is with airline} j \\ 0 & \text{otherwise} \end{cases}$$

$j \in$ {American Airlines Network, Delta Air Line Network, Jetblue Airways, Southwest Airlines, Spirit Airlines, United Airlines Network}

$Time_r$ represents the time of day the flight occurs. Most indices represents an hour starting with 1 indicating the time 6am - 7am, except 18 which indicates 23hr - 6hr

$$Time_r = \begin{cases} 1 & \text{if flight occurs at time} r \\ 0 & \text{for all other times} \end{cases}$$
$$r \in \{1, 2, ..., 18\}$$

$$e_{i,j} = \text{whether you pick the a flight on airline j leaving from airport i.}$$
$$schedule_{i,r} = \text{whether you are taking the flight from airport i at time r}$$

9.3 Objective Function

$$S_l * \sum (e_{i,j} * [1 - OnTime_{i,j}]) \tag{1}$$
$$S_l * \sum (schedule_{i,r} * [1 - Timely_{i,r}]) \tag{2}$$
$$S_c * \sum (e_{i,j} * Cancelled_{i,j}) \tag{3}$$
$$S_d * \sum (e_{i,j} * Diverted_{i,j}) \tag{4}$$

minimize ① + ② + ③ + ④

9.4 Constraints

$$e_{i,j} \leq airport_i \tag{5}$$
$$e_{i,j} \leq airline_j \tag{6}$$
$$e_{i,j} \geq airport_i + airline_j - 1 \tag{7}$$
$$0 \leq e_{i,j} \leq 1 \tag{8}$$
$$schedule_{i,r} \leq airport_i \tag{9}$$
$$schedule_{i,r} \leq time_r \tag{10}$$
$$schedule_{i,r} \geq airport_i + time_r - 1 \tag{11}$$
$$0 \leq schedulei, r \leq 1 \tag{12}$$
$$\sum airport_i = 1 \tag{13}$$
$$\sum airline_j = 1 \tag{14}$$
$$\sum time_r = 1 \tag{15}$$

10 Explanation of the Model

Given a trip from San Francisco to New York, there are 3 major decisions that influence the probability of stressful events occurring: which New York airport you fly to, which airline you fly on, and what time your flight is. The stressful events we incorporated in this model are: the flight being cancelled, diverted or late. Constraints number 13, 14 and 15 insist that only one airport and airline will be used at a certain time; this mimics the choices needed to take an one-way flight.

Each stressor makes up a section of the objective function. Functions 1, 3, and 4 determines the expected amount of stress from each stressor that would occur if the chosen flight is on airline j leaving from airport i; this would happen when the variable $e_{i,j} = 1$. The stressor for objective function 1 is flight being late, the stressor for 3 is the flight being cancelled and 4 is the stressor for being diverted. We also wanted to be able to capture the likelihood of being delayed based upon the flight's departure time. This had to be a separate variable since the data wasn't broken down by airport and airline but rather airport and time of day. The impact that the time of day has with a flight being delayed is calculated through objective function 2.

The variable $e_{i,j}$ is needed since we cannot multiply the binary variables, airport i and airline j, together in the objective function while still keeping the formulation linear. Constraints 5 through 8 are used to enforce $e_{i,j}$ to be either be 0 or 1 based upon the airport and airline decision. If airport i and airline j are both one for some i and j, then we want $e_{i,j}$ to be equal to 1. Similarly $schedule_{i,r}$ was created utilizing constraints 9 through 12.

11 Explanation of the Code

The attached python file Project.ipynb implements the model based on the formulation above and leverages the Gurobi package to solve it. First, the program must read in the probability of stressors from the dataset used in the model (found in the attached Data.xlsx file), which is split up based on airline, airport, and flight time. Then, the python program offers the user the opportunity to input their own custom stressor weights, which is explained as how much stress on a scale of 1 to 10 each events would cause them if it occurred. If the user chooses not to input their own values, default values for the stressors weights are used, which were generated from our research and experience. The estimated weights are a 10 for flight cancellation, a 2 for a late flight, and a 5 for a diverted flight. From there, the program formulates the model using the Gurobi package, adding constraints and defining the objective function as described in the formulation. After optimizing this, the program outputs a csv file containing the name of the airport, the name of the airline, the time of the flight, and the objective value for the optimal solution.

12 Optimization Results

The project_sol.csv file attached is the output csv file generated by running the program with the default stressor weights. The solution selects one single option for each decision, with JFK as the airport, Delta as the airline, and a flight scheduled from 6am to 7am. The objective value in this model is 0.266, which has little practical meaning, but can be interpreted as the "expected stress level" of the travel plan which we are seeking to minimize. We note that since each probability is above zero and exactly one of each of the binary variables for airport, airline, and flight time will be equal to one, that the objective value is greater or equal to zero, or lower bounded by zero. However, it is expected that the objective value would be close to zero, since the stressors considered in this model have relatively low probabilities, with less than 1% chance for diverted or cancelled flights, and around a 15-30% chance of a late flight and stressor weights are between 1 and 10. We also modeled the most stressful flight which was determined to be the airline Southwest from the airport, EWR between the times 23:00 - 24:00.

13 Result Implications

Some of the interesting trends that could be analyzed include the correlation between how much you spend on flights or the departure times, and how likely stressors are to occur. The timings of the flights varied between the most stressful and least stressful flights; the least stressful flight occurred during the morning comparatively the most stressful at night. It is not surprising since morning flights aren't impacted by a backlog of other flights; for example a flight leaving at 23:00 may depend on a previous flight arriving on time since they would use the same plane. Delta showed up in our optimal solution and is the most expensive airline to run so there may exist a correlation between expensive flights and less stressful ones. Similarly when we maximized stress Southwest one of the cheapest airlines to run also was a part of our most stressful flight to take. The average fare costs from the three different airports didn't vary significantly though they are some of the most expensive in the United States. (Mazareanu, 2019) One could seek to optimize both cost and stress by incorporating cost as a factor of stress. It's important to consider the cost as some consumers are affected by the price elasticity of airline fares.

14 Extensions of the Model

14.1 Added Stressors

One of the primary ways to expand this model would be simply to add more stressors to the model. Within the dataset used, there is data on other stressful events such as lost luggage, involuntary denial to board, and many more. However, these stressors required more manipulation to fit into the model, and it was decided that the 3 stressors chosen would sufficiently exhibit the potential value of the model. The level of stress associated with each added stressor would similarly have to be quantified by weights and then multiplied by the probability of the event for the chosen airline, airport, and flight time. Essentially, they would follow the same pattern as the stressors currently in the model, but would require custom constraints and variables added depending on which decision variables affect them.

14.2 Added Decisions

The decisions we chose to represent the model were airport, airline, and flight time, but those are not the only decisions one makes when planning travel, and not the only ones that affect your level of stress or probability of stressors occurring.

Another added decision could be once again deciding airline, airport, and time for the return flight back to San Francisco. While this is a practical addition to the model, as travelers often want to return back home from a vacation, it essentially duplicates the decisions and constraints formulated for the flight to your destination. It would simply incorporate the data on airlines, airports, and flight times for Y to X instead of X to Y. Thus, this expansion on the model was omitted, however the model formulated would simply need to reference the corresponding dataset to achieve this expansion.

You could also further expand the model by adding in more options within the pre-existing decisions from the dataset, so you could add more airlines and airports to choose from. You could even expand the scope of these decisions by allowing variable destinations, such as flying from Toronto to Los Angeles, add different times of the year to fly, or the ability to take connecting flights to your chosen destination.

14.3 Stress Alleviators

Another expansion to the model could be stress alleviators, such as compensation for inconvenience, complimentary food or beverage on the flight, or other forms of satisfaction guaranteed. This would create a model by which the objective function balances the likelihood of events that reduce stress and those that increase stress given different travel plan options.

The expansions discussed are only some of many possible additions which could be included to enhance the accuracy and scope of the model. This showcases the vast potential of this optimization model, and just how deeply it can seek to promote a happier, stress-free travel experience.

15 Conclusion

Overall, the aim of the report was to derive a process whereby the complexities of stressful travel can be quantified such that they can be analyzed and therefore reduced. This project is a demonstration that when optimization theory, high quality data, and creativity is combined, there is a world of possibility in the kinds of problems you can seek to solve.

16 Bibliography

"Airline On-Time Statistics." *Bureau of Transportation Statistics*, United States Department of Transportation, 17 Dec. 2019, www.bts.gov/newsroom/air-travel-consumer-report-october-2019-numbers.

Groves, William, and Maria Gini. "On Optimizing Airline Ticket Purchase Timing." *ACM Transactions on Intelligent Systems and Technology*, vol. 7, no. 1, 2015, pp. 1–28., doi:10.1145/2733384.

Ivancevich, John, et al. "Business Travel Stress: a Model, Propositions and Managerial Implications." *Work Stress*, vol. 17, no. 2, 2003, doi:10.1080/713868351.

Kiseleva, Julia, et al. "Where to Go on Your Next Trip?" *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 3 June 2015, doi:10.1145/2766462.2776777.

Li, Jun, et al. "Are Consumers Strategic? Structural Estimation from the Air-Travel Industry." *Management Science*, vol. 60, no. 9, Sept. 2014, doi:10.2139/ssrn.1939033.

Madden, Duncan. "Revealed: The Most Stressful Aspects Of Airport Travel." *Forbes*, Forbes Magazine, 17 Sept. 2019, www.forbes.com/sites/duncanmadden/2019/09/17/revealed-the-most-stressful-aspects-of-airport-travel/57f57584cab7.

Mazareanu, E. "U.S. Airlines: Domestic Cost per Available Seat Mile 2018." *Statista*, 24 Apr. 2019, www.statista.com/statistics/527881/us-domestic-cost-per-available-seat-mile-by-airline/.

Shanahan, Matthew J. "Mathematical Modeling of Stress Management via Decisional Control." *Western Libraries*, University of Western Ontario, 14 Dec. 2015.

Yu, Gang, and Jian Yang. "Optimization Applications in the Airline Industry." *Handbook of Combinatorial Optimization*, Jan. 1998, pp. 1381–1472., doi:10.1007/978-1-4613-0303-$9_2$3.