Robust Optimization of Travel Management in Europe

Mason Chen

***Abstract*—The objective was to manage European travel plans by limiting the time to 28.8 hours and cost to $500 used in transportation. Design constraints were implemented, ensuring that the model was practical. There are 5 travel segments which have 2 choices: flight/train. Taking a train would better when traveling <500km while taking a flight is preferred when traveling >1,000km. Research was further done on the aerodynamics involved in helping an airplane fly such as lift and drag and physics involved in trains such as friction and acceleration. Based on the costs & durations, only the Paris-Munich route is debatable on which transportation to use. A multiple regression model was built on mean time & expense. The predictive DOE model shows that Paris-Munich by train is more desirable than flight. Although the neural model was conducted, it was not able to outperform the structured DOE model in terms of desirability. However, we should not exclude the neural method as the Structured DOE and the modern AI Algorithm should coexist. The robust design Monte Carlo simulation shows that there is an 8% risk of not meeting the $500 budget. By raising the importance of time, the new optimal route can meet requirements.**

***Index Terms*—Travel, optimization, Europe, DOE, Monte Carlo simulation.**

I. INTRODUCTION

Although many people like traveling, there are always constraints that might inhibit traveling such as not having enough time, limited budget, and poor travel quality. The objectives of this paper are to manage the travel plan in Europe, build a statistical and neural model to predict travel duration and expense, and conduct a Robust Design to optimize the travel plan. A great way to save time and money when traveling across Europe is by taking the trains. Taking Eurostar trains is a more convenient and economical choice than taking flights. However, Eurostar only has direct trains among major cities such as London, Paris, Brussel, Lyon, and Avignon. Taking trains among other smaller cities in Europe may not be a better choice than taking a flight. This paper designs a summer travel package on visiting five major European cities: London, Paris, Amsterdam, Prague, and Munich. The start and end destination will be at the Paris Charles de Gaulle (CDG) Airport which is applicable for most visitors. CDG airport is the largest international airport in France and the second largest in Europe. This paper will only consider the traveling time and expense of taking either train and/or flight among five destinations. After optimizing the travel plan, this project will also consider the lodging expenses. Which has the longest total time – flight or train? Although flight is mostly faster in terms of transportation time, if taking into consideration the lengthy check-in time, the total flight time may be greater than the total train time. However, the train may not always take the most direct route as it has to be on land. After comparing the train and flight distances, a multiple of 1.3x is the average factor of train compared to flight. If the flying distance is around 500 kilometers, the flying time is about 1.25 hours plus a 3 hours check-in and check-out time, a total of 4.25 hours. If train is taken, the train's distance will be 500 kilometers times 1.3 which equals 650 kilometers. The train transportation time is close to 2.5 hours plus an hour of check-in and check-out time, a total of 3.5 hours. When the flight distance is under 500 kilometers, we can safely conclude that train is a better choice than flight. If the flying distance is around 1000 kilometers, the flying time is around 2 hours plus 3 hours check-in and check-out time, a total of 5 hours. If train is taken, the train's distance will be 1000 kilometers times 1.3 which equals 1300 kilometers. The train transportation time is close to 5 hours plus an hour of check-in and check-out time, a total of 6 hours. When the flight distance is over 1000 kilometers, we can safely conclude that flight is a better choice than train. The breakeven point is between 500-1000 kilometers, which is the distance that most major cities lie in.

II. DEFINE PROJECT SCOPE

In the previous section, we explored the science behind flight and train and concluded that we needed an optimal model to predict whether train or flight is better.

A. Alternate Transportation Choices

In Fig. 1 [1], the shortest route for train is 3123 kilometers while by flight is 2359 kilometers. To simplify this project, we will only consider the two shortest (green) routes – clockwise and counterclockwise. The first route is Paris-London-Amsterdam-Prague-Munich-Paris and the second is Paris-Munich-Prague-Amsterdam-London-Paris. Considering both routes separately is important as the ticket price and flying duration one way may be different in opposite directions.

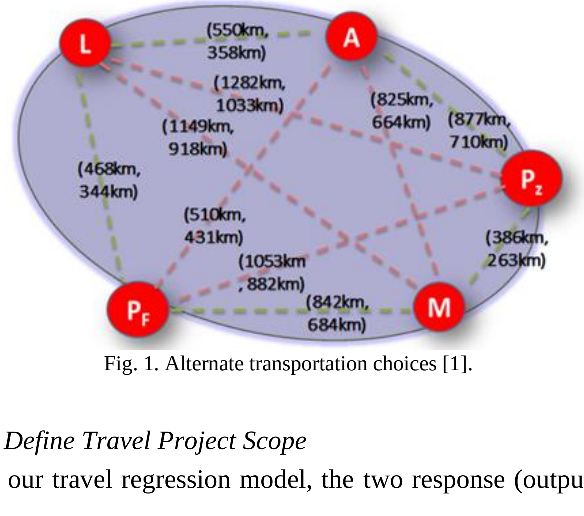


Fig. 1. Alternate transportation choices [1].

B. Define Travel Project Scope

In our travel regression model, the two response (output)

Manuscript received December 2019; revised May 31, 2020.

Mason Chen was with Stanford OHS, USA (e-mail: mason05@ohs.stanford.edu).

doi: 10.18178/ijimt.2021.12.1.902

variables are the total intra-city transportation time and the total intra-city transportation expense. The three input variables are travel route (sequence), each intra-city transportation time, and each intra-city transportation expense. In addition, to simplify the regression model, several design constraints would be considered throughout this project. We will collect all transportation raw data on July 11, 2018, and book tickets by May 26, 2018, start and end at Paris CDG Airport or the city's main station, add 3 hours check-in and check-out time for flight, and 1 hour check-in and check-out time for Eurostar, take Economy flight seat or 2nd Class train seat (all one-way ticket), only consider direct flight or direct train if available; Otherwise, take one stop in transition, only consider flight or train after 9 AM and arrive by 9 AM within the same day, will not consider driving or ferry transportation among five cities, and won't consider flight/train delay factor.

C. *Set Travel Transportation Requirements*

To manage this special travel package among five major cities in Europe, several reasonable travel requirements are set. The total intra-city transportation duration must be less than 28.8 hours (12 days of travel, 9 AM to 9 PM, and 12 hours * 12 days = 144 hours total). Less than 20% of the 144 hours is allocated to intra-city transportation time (< 28.8 hours). Secondly, the total intra-city transportation expense cannot exceed 500 USD (12 hotels = 2,400 USD (200 USD/night), meals 900 USD (75/day), other tour/local transportation (1,200 USD = 100/day), total intra-city transportation expense less than 500 USD).

III. ESTABLISH PREDICTIVE METHODOLOGY

In Section III, a design of experiment (DOE) methodology was utilized to determine the optimal route. A structured Doe was first designed before conducting a robust design – Monte Carlo Simulation.

A. *Design a Structured DOE*

A special JMP Definitive Screen Design (DSD) was conducted in order to optimize the transportation route. A definitive screening design (DSD) would be conducted to optimize the neural algorithm. Areas, where definitive screening designs are superior to standard screening designs, include identifying the causes of nonlinear effects by fielding each continuous factor at three levels and avoiding confounding between any effects up through the second order [2], [3]. There are five one-way transportation segments of categorical input variables (flight or train), a total of 18 DSD runs. To ensure the DSD structure, three examination criteria was done before conducting the DSD simulation runs on the Neural Network algorithm [4], [5]. The first power analysis checked whether the DSD run size was sufficient. If the run size was too small, the 95% confidence interval of any effect term will be very wide. Then, the power level would indicate the probability of the predicted sign is still valid. In Figure 10, all power levels are above 88% (small run size concern). The second confounding color-map analysis is to investigate whether any Resolution I or Resolution II confounding concerns between any main effect. The confounding severity is indicated by the color map (from 0% correlation – blue to 100% correlation – red). The diagonal is always in red color. There is very mild Resolution II confounding (correlation = 0.33) due to categorical variables. Therefore, no severe confounding concern was noticed.

B. *DSD Optimization Result*

Section III.B will provide the DSD results of Section 3.1 DSD execution. The objective of this DSD is to demonstrate the optimal transportation route in order to minimize two intra-city transportation goals: total duration and total expense. Among the five transportation segments, Amsterdam to Prague segment has shown the biggest impact to both expense and duration responses according to the DOE model as shown in Fig. 2 [6]. Taking flight is a much better choice than taking the train since there is no Eurostar train service from Amsterdam to Prague. There is also no direct train available between the two cities. This train route is not very popular and the train ticket is way more expensive than taking a flight. Next, there are two competing patterns between two responses on (1) Prague to Munich, and (2) London to Amsterdam. Prague to Munich segment has more impact on the expense and taking direct train would significantly reduce the expense around $100 USD as compared to taking a flight. Instead, for the London-Amsterdam segment, taking a flight can shorten the transportation duration time by more than 200 minutes. The driving/flying distance ratio of the London-Amsterdam route is 550km/358km ~ 1.54 which is higher than the typical 1.3 ratio in Europe. From London-Amsterdam, the train needs to pass Paris first which means the flight can be more direct (as it can fly over waters while a train cannot). The fourth sensitive segment Munich to Paris also favors train choice in order to meet the expense requirement more. The last segment Paris to London favors train to meet the duration requirement more. This explains why most visitors would rather take Eurostar train across the English Strait. The optimal design can achieve the expected expense at 392 USD (below 500 USD) and duration 1,385 minutes (below 2,880 minutes requirement). However, the overall optimal design can only meet both requirements at 65% desirability and still does not include noise factors.

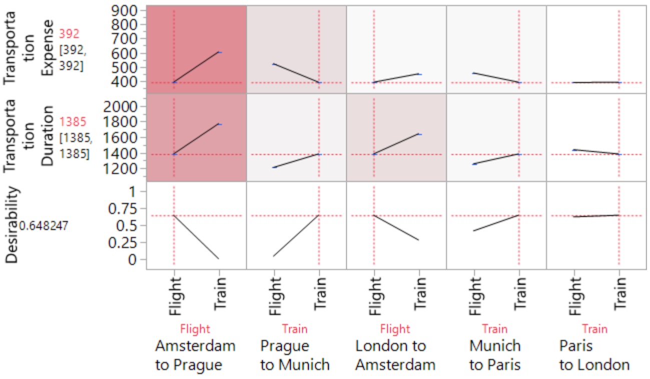
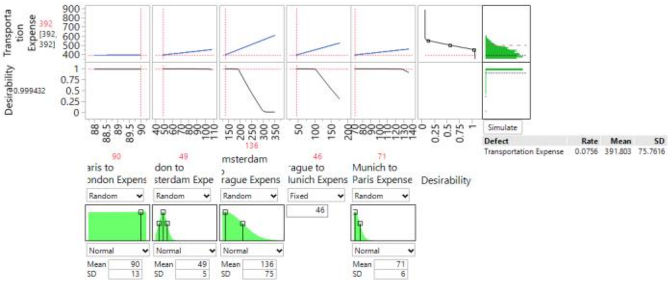


Fig. 2. Profiler sensitivity and optimization.

C. *Monte Carlo Simulation*

Section III.B optimal route design would not consider the noise factors such as duration and ticket price distribution. Even within the same day at the same station, the transportation duration and the ticket price may be different from peak hours to off-peak hours. In order to accurately estimate the total transportation duration and expense, considering all the uncertain noise factors is desperately needed. JMP Profiler Monte Carlo simulator is very powerful to simulate these random noise factors. Monte Carlo simulation, or probability simulation, is a technique used to understand the impact of risk and uncertainty in finance, project management, cost, and other forecasting models. Their essential idea is using randomness to solve problems that might be deterministic in principle. They are often used in physical and mathematical problems and are most useful when it is difficult or impossible to use other approaches. Monte Carlo methods are mainly used in three problem classes: [7] optimization, numerical integration, and generating draws from a probability distribution. Monte Carlo Simulation helps to discover the distribution of model outputs as a function of the random variation in the factors and model noise. The simulator in the profilers provides a way to set up the random inputs and run the simulations, producing an output table of simulated values. As shown in Fig. 3, JMP Profiler Simulators use Normal Distribution to simulate the two responses duration and expense) by including the variability of flight duration and ticket price.



Defect	Rate	Mean	SD
Transportation Expense	0.0756	391.803	75.7616

	Mean	SD
aris to ondon Expens	90	13
don to sterdam Expe	49	5
msterdam) rague Expens	136	75
Munich to Paris Expense	71	6

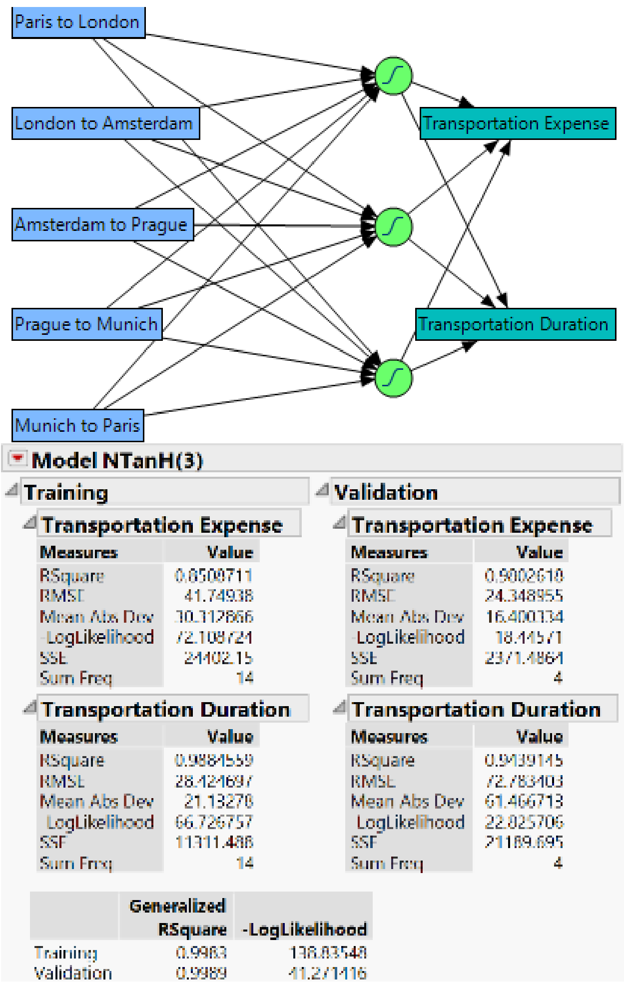
Fig. 3. Monte Carlo simulation of transportation expense.

The distribution standard deviation was determined by calculating the real duration/ticket price data. The two requirements are also input variables for the upper specific limit (USL) for the JMP Profiler to estimate the non-conforming defect rate. The Monte Carlo Simulation results have shown 0% probability of not meeting the 28.8hours duration requirement and 8% probability of not meeting 500 USD transportation budget. Even though the predicted expense is 392 USD (seems enough buffer from 500 USD), when considering the noise ticket price distribution, there is still an 8% risk that the total transportation expense may exceed 500 USD. In order to avoid this 8% uncertainty, booking the flights/train tickets – that are not scheduled during the peak hours – earlier can help eliminate this risk.

D. Neural Network Modelling

The neural algorithm uses the TanH transformation and the black box scenario as shown in Fig. 4a. In Figure 4b, the neural algorithm can achieve >90% r-square on both the training set and validation set. No overfit concern was observed in the neural analysis.

There is a similar sensitivity pattern (seen in the red color) between the two optimization algorithms. The DSD RSM model has shown slightly higher desirability than the neural model as shown in Fig. 5. Even though the DSD approach outperformed the modern neural AI method, they both should coexist. The AI method can detect hidden patterns while the structured DOE method can be conducted to validate the AI pattern more reliably.



Model NTanH(3)

Training

Transportation Expense

Measures	Value
RSquare	0.8509711
RMSE	41.74938
Mean Abs Dev	30.312866
-LogLikelihood	72.108724
SSE	24402.15
Sum Freq	14

Transportation Duration

Measures	Value
RSquare	0.9884559
RMSE	28.424697
Mean Abs Dev	21.13270
-LogLikelihood	66.726757
SSE	11311.488
Sum Freq	14

Validation

Transportation Expense

Measures	Value
RSquare	0.9002610
RMSE	24.348955
Mean Abs Dev	16.400334
-LogLikelihood	18.44571
SSE	2371.4864
Sum Freq	4

Transportation Duration

Measures	Value
RSquare	0.9139145
RMSE	72.783403
Mean Abs Dev	61.466713
-LogLikelihood	22.025706
SSE	21189.695
Sum Freq	4

	Generalized RSquare	-LogLikelihood
Training	0.9983	138.83548
Validation	0.9989	41.271416

Fig. 4. (a) Neural model outline (b) TanH transformation statistics.

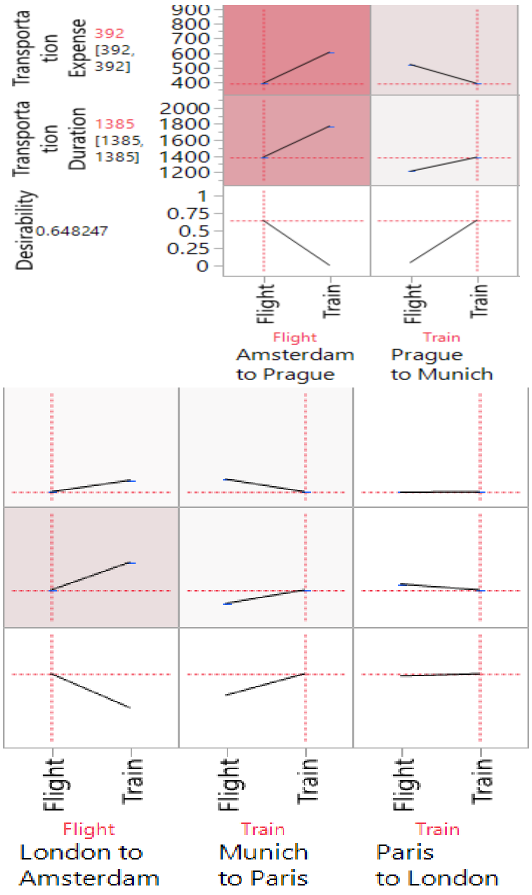
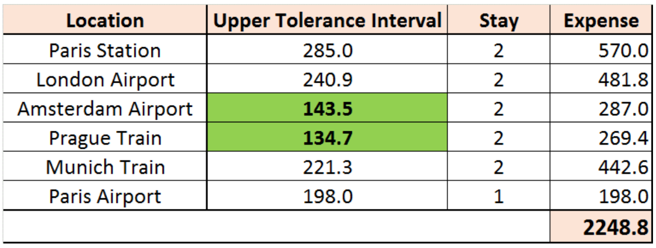


Fig. 5. DSD method.

E. Include Lodging Expense Model

In the end, the optimal route is Paris (train) – London (flight) – Amsterdam (flight) – Prague (train) – Munich (train) – Paris (flight). The budget is less than 2,400 USD for the twelve nights among the 5 cities with at least two nights at each stop except for the last stop in Paris. An extra one night has not been determined, depending on which location has a cheaper lodging. To consider the lodging expense, the criteria used were to choose a 3-star hotel within 3 miles of the airport or train station with a guest rating above 9 and book the hotel 6 months in advance. A 1-sided upper tolerance interval was implemented to determine the worst scenario lodging budget lodging expense. As shown in Figure 6, the total budget expense is less than 2,400 USD with 11 fixed/decided nights, and one remaining lodging budget (excess of 151 USD). It was decided to stay one more night at Prague to save the 16 USD transportation shortage (if wanting to stay at another city for the last night).



Location	Upper Tolerance Interval	Stay	Expense
Paris Station	285.0	2	570.0
London Airport	240.9	2	481.8
Amsterdam Airport	**143.5**	2	287.0
Prague Train	**134.7**	2	269.4
Munich Train	221.3	2	442.6
Paris Airport	198.0	1	198.0
			2248.8

Fig. 6. 1-sided upper tolerance interval.

IV. CONCLUSION

This paper has demonstrated an effective methodology of managing Travel in Europe by studying the Transportation Systems in Europe. Physics was applied on the friction and acceleration of the train as well as aerodynamics and forces on the airplane. A structured DOE and neural algorithm were designed and built predictive models of minimizing both the travel duration and expense. Although the DOE had a higher desirability than the neural model, the modern and classical models should coexist. Monte Carlo simulation method was conducted by considering the variability of random flight/train duration and expense. Although there is an observed 8% risk probability of not meeting the transportation expense budget, it can be solved by booking tickets earlier. The same methodology can be applied to travel in China, Japan, Taiwan where the high-speed train systems are well established.

CONFLICT OF INTEREST

No conflict of interest.

ACKNOWLEDGMENT

The author acknowledges Dr. Charles Chen for his support.

REFERENCES

[1] M. Chen and C. Chen, "Apply "STEAMS" methodology on managing Europe travel," *ASA JSM Proceedings*, pp. 946-958, 2019.

[2] B. Jones and C. J. Nachtsheim, "Definitive screening designs with added two-level categorical factors," *Journal of Quality Technology*, vol. 45, pp. 121-129, 2013.

[3] B. Jones and C. J. Nachtsheim, "Blocking schemes for definitive screening designs," *Technometrics*, vol. 58, no. 1, pp. 74-83, 2016.

[4] A. Miller and R. R. Sitter, "Using folded-over nonorthogonal designs," *Technometrics*, vol. 47, no. 4, pp. 502-513, 2005.

[5] A. Errore, B. Jones, W. Li, and C. Nachtsheim, "Using definitive screening designs to identify active first- and second-order factor effects," *Forthcoming, Journal of Quality Technology*, 2016.

[6] H. L. Anderson, "Metropolis, monte carlo and the MANIAC," *Los Alamos Science*, vol. 14, PP. 96-108, 1986.

[7] Mason, C, "Robust optimization of travel management in Europe," *ASA SDSS Proceedings.* PP. 959-967, 2019.

Mason Chen is currently a student at Stanford OHS and serves as the student ambassador and webmaster for STEAMS. Having started STEAMS since its inception in 2014, he has held various roles such as president of the student chapter from 2017 to 2019. Through STEAMS, he has published more than 20 conference proceeding papers as first, second, or third author. As first author, he has won numerous awards including the best conference proceeding paper award in the 2018 JMP discovery summit as well as finishing 1st place three times for the STEM presentation competition at IEOM conferences. He has also certified the IBM SPSS Statistics Level I, II, Modeler Level I, and IASSC Yellow Belt, Green Belt, and Black Belt.