SURROGATE MODELING OF AIRCRAFT PERFORMANCE

SYED MUHAMMAD TASHFEEN JAVAD

NUST2015-MS SYS 00000148742

Supervised by

Dr. Adnan Maqsood

RESEARCH CENTRE FOR MODELING AND SIMULATION

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY

(NUST), ISLAMABAD

September 2017

SURROGATE MODELING OF AIRCRAFT PERFORMANCE

SYED MUHAMMAD TASHFEEN JAVAD

Research Centre for Modeling and Simulation

A thesis submitted to the National University of Sciences & Technology in

partial fulfillment of the requirement for the degree of

Masters of Science

In

Systems Engineering

2017

Annex A to NUST Letter No. 0972/102/Exams/Thesis-Cert dated 23 Dec 16.

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis written by Mr/Ms **Syed Muhammad Tashfeen Javad_** Registration No. _**NUST2015-MSSYS0000148742**_ of **__RCMS___** has been vetted by undersigned, found complete in all aspects as per NUST Statutes/Regulations, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature with stamp: ___________________________

Name of Supervisor: ____________________________

Date: ____

Signature of HoD with stamp: _____________________

Date: ____

Countersign by

Signature (Dean/Principal): ________________

Date: __________________________________

STATEMENT OF ORIGINALITY

I hereby certify that the work embodied in this thesis is the result of original research and has not been submitted for a higher degree to any other University or Institution.

____
Date

____
S M Tashfeen Javad

Dedication

I dedicate this effort to all those who have assisted me in any possible way to become what I am today. Their sacrifices seeded my success especially my parents who showed their devoted attention and to faculty members who inspired me all the way.

Acknowledgements

I am thankful to Allah Almighty for bestowing me with the courage, knowledge and health to carry out this thesis and to my Parents and family members, without their support this thesis was impossible.

Special thanks to my supervisor Dr. Adnan Maqsood for his support, motivation and keen interest in this work. Without his support, I would not have been able to complete my thesis.

I would also express heartfelt thanks to my thesis committee members Dr. Zamir Hussain for his assistance in statistical area of my research and Dr. Ammar Mushtaq for their skillful assistance and guidance.

I am also thankful to staff members and colleagues of RCMS who helped me through the difficulties I faced during the thesis.

Abstract

The aviation industry has expanded enormously during last few decades. The advent of new technologies has made the aircraft design process more complicated. There are certain modeling techniques that can help circumvent the existing design process and yet yield fair approximates. One such approach that helps us to arrive at the outcome with reasonably good approximation is known as surrogate modeling. Surrogate models help circumventing complicated analytical methods, time-consuming simulations and expensive experimental techniques typically used during the design process.

For the said purpose, a database of aircraft including military jets, commercial airliners and unmanned air vehicles was developed from commercially available data. Data collected was scrutinized into dependent and independent variables. Aircraft performance parameters needed to be estimated so they were dependent variables and aircraft geometric parameters being predictors were independent variables. Scalability trends using power laws were developed between dependent and independent variables. The scalability study provided the initial design bounds and developed the foundations for further developing of surrogate models. As aircraft design is a complex process and a single variable might not be sufficient approximations for all scenarios, therefore, surrogate models using multiple linear regression technique were developed. The developed models estimated the dependent variables with high confidence. Moreover these models were validated using a three step process which included verification of each model using quantitative criteria, comparing the models with analytical equations and checking the prediction accuracy of each model. It is thus validated that a multiple linear regression models are best fitted for modeling surrogate models for aircraft performance parameters. The approach proposed in this thesis holds a strong utilization potential among the aircraft design community and aero modelers during preliminary design phase.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

b	wing span
C_L	coefficient of Lift
C_D	coefficient of Drag
C_{D0}	coefficient of Drag at zero lift
L/D	lift to drag ratio
MAE	mean absolute error
MLR	multiple linear regression
MALE	medium altitude long endurance
MUAV	mini unmanned aerial vehicle
NN	neural network
OEM	original equipment manufacturer
PCA	principle component analysis
P_A	Power available
RSM	response surface method
RBF	radial base function
RMSE	root mean square error
ROC, R/C	rate of climb
S	wing surface area
SFC	specific fuel consumption
T_A	Thrust available
TSFC	thrust specific fuel consumption
UCAV	unmanned combat aerial vehicle
VFM	variable fidelity model
V_{max}	maximum velocity
V_∞	free stream velocity
W, MTOW	maximum takeoff weight

CHAPTER 1

INTRODUCTION

Scalability trends can be observed in every living organism. Insects, birds and animals all have a scalable relationship among their ages, metabolic rates, speeds and so on [1]. Similarly the life at cities follows scalability laws [2]. Inspired by these relationships, scalability studies have been conducted in the past for advancement in technology. Semiconductor devices and electronics have demonstrated scalable growth. Gordon E. Moore [3] predicted that the number of transistors in dense electronic circuits will double approximately every two years. The prediction remained accurate for several decades. The traditional approaches to model complex systems are usually costly, resource exhaustive and time intensive. Another interesting application of predictive analysis was seen in fuel consumption modeling of an aircraft using flight operations quality assurance data [4]. These analyses have proved that certain models can be developed which will circumvent the existing design process and yet yield results acceptable to the designers. Such models that help us to arrive at the outcome with good approximation are known as surrogate models.

The primary focus of this research is to develop an alternate procedure than existing analytical / experimental techniques in estimating aircraft performance during design phase. The conventional design process include defining detailed geometric descriptions, estimating aerodynamic data from rigorous computational / analytical / experimental techniques, followed by aircraft performance modeling using point mass models. The whole process sometimes gets tedious especially when the design activity is undergoing several iterations. The process is cumbersome most of the times and the results achieved in the end are still approximates. The designers over decades have come up with different designs following same fundamental abstract principles. We have now sufficient amount of data for different aircraft for which we can generate approximate / surrogate models from earlier designs. The initial design process is always an iterative process and requires validation from analytical / computational methods, before the design can go into the experimental phase. The solution needs to be well optimized before the sketch goes into prototyping, as the financial blows from a failed prototype are enough to shelf the whole project. The aviation history is witness to the fact that flaws which were overlooked in the design phase and ended up in prototyping, were failed thus jeopardizing the project. The solution to this problem lies in optimizing the initial design in an efficient manner so that later details can be studied and looked with utmost care. Surrogate models are a solution to such problems and can act as predictors for futuristic designs. Although the results of using such models might be crude but the time saving outstrip this limitation. Moreover, these models can help novice designers to generate different estimates without fundamental knowledge.

1.1 MOTIVATION

The statistical techniques have been used in almost every field and have produced useful results. The estimates deduced from statistical approach are very much comparable with analytical solutions and thus can be a benefiting tool to build surrogate models. The uses of these techniques are seen from predicting human behavior to estimating machine performance. Thus these diverse studies conducted using statistical approach has been the source of motivation to conduct a statistical based study in the field of aircraft performance estimates.

1.2 INTRODUCTION TO CURRENT RESEARCH

Surrogate models will be used to predict aircraft performance parameters using regression techniques. The results from these surrogate models will then be compared with the existing analytical solutions. It is to be highlighted that aircraft performance parameters calculation require certain data that can only be calculated using expensive experimental technique such as wind tunnel testing. Surrogate models will provide a crisp and effective solution for the design engineers in initial design process where using basic geometric data, one can calculate the aircraft performance and use it as a starting point.

1.2.1 RESEARCH OBJECTIVES

The scope of research is to estimate aircraft performance parameters (range, maximum velocity, endurance and rate of climb) using aircraft geometric and general parameters. The aircraft data will include different platforms so as to make the models more inclusive. Following are the objectives of current research:

(a) To build surrogate models for aircraft performance parameters which include ceiling, range, maximum velocity, endurance and rate of climb

(b) To assess the suitability of the proposed models using statistical and experimental analysis

In order to achieve the above mentioned objectives following sequence will be followed

a) Data collection

i. Collection of different type of aircraft's parameters

ii. Data include different aircraft categories

b) Application of statistical technique

i. Use of power law's to study scalability trends

ii. Using regression technique to develop models

iii. Validation of both techniques to develop surrogate models

c) Comparison of results

i. Comparison of Statistical results with Analytical equations of aircraft parameters and evaluate the prediction accuracy of surrogate models

1.3 POTENTIAL CONTRIBUTION

The valuable potential contributions made through this scholarly work will be:

- **i.** Assessment of surrogate modeling application in an aviation design field
- **ii.** Identification of exact technique as to how these surrogate models will be developed
- **iii.** Validation of mentioned technique using certain validation methods

1.4 THESIS ORGANIZATION

This thesis is organized as follows.

Chapter 1- Introduction

This chapter gives a background on fundamental challenges in aircraft design process and why there is a need to build such models that can help the designers in initial design phase.

Chapter 2-Literature review

This chapter gives a detail review of use of surrogate modeling as an instrumental tool for designers. The relevant work in the field has also been discussed in detail.

Chapter 3- Methodology and Problem formulation

The chapter explains the detailed methodology employed to circumvent the existing tedious design process.

Chapter 4- Single variable surrogate modeling

Single variable models were developed using power laws

Chapter 5- Multiple variable surrogate modeling

Multiple variable models are developed using multiple linear regression (MLR)

Chapter 6- Model accuracy and validation

Both models are then compared for accuracy and further validation for the selection of fitted model was conducted.

Chapter 7- Conclusion and future work

The thesis concludes in where recommendations are suggested for future work.

CHAPTER 2

LITERATURE REVIEW

Surrogate modeling has seen much application in diverse fields. Their contribution in engineering fields is worth appreciating. The models have helped engineers to tailor the existing procedures and come up with efficient designs and prototype thereby saving time and resources. Literature review can be divided onto two main areas, scalability study and complex system modeling using different techniques. Surrogate models can be broadly used in two ways to benefit the existing systems:

i. **Design estimations**: Estimating the outcome by building certain models that can act as a quick alternative to existing methods. This estimation helps in cutting the modeling and simulation computational time and cost.

ii. **Design optimization**: Improving the design efficiency by finding the global or local optima rather than employing the time exhaustive conventional optimization techniques.

2.1 Scalability research literature

Most recent and noticeable scalability studies are conducted by Geoffrey West [1, 2]. These studies contain relationships in almost every field of life ranging from animals metabolic rates to the life in big cities of the world. These studies span around the simple power law relationships i.e. $Y = \alpha * X^{\beta}$ with the exponents that are simple multiples of ¼ (e.g. ¼, ¾ etc.). The graph plotted among the masses and metabolic rates follows a linear relationship as shown in Figure 1. It explains that as the body masses are increased, metabolic rates needs to be increased for the body to perform its intended tasks. If the metabolic rate is not inconsistent with body mass, it means that body is under certain disease attack. Scaling features taken from data analysis for animals (mammals) include, volume of blood scales with the mass (M), $V_b \propto M^{1.02}$, size of heart scales with the mass, $M_h \propto M^{0.98}$, frequency of heartbeat scales as: $f_h \propto M^{-1/4}$and volume of lungs scales with mass: $V_l \propto M^{1.04}$ and The tidal volume of lungs scales as: $V_t \propto M^{1.04}$.

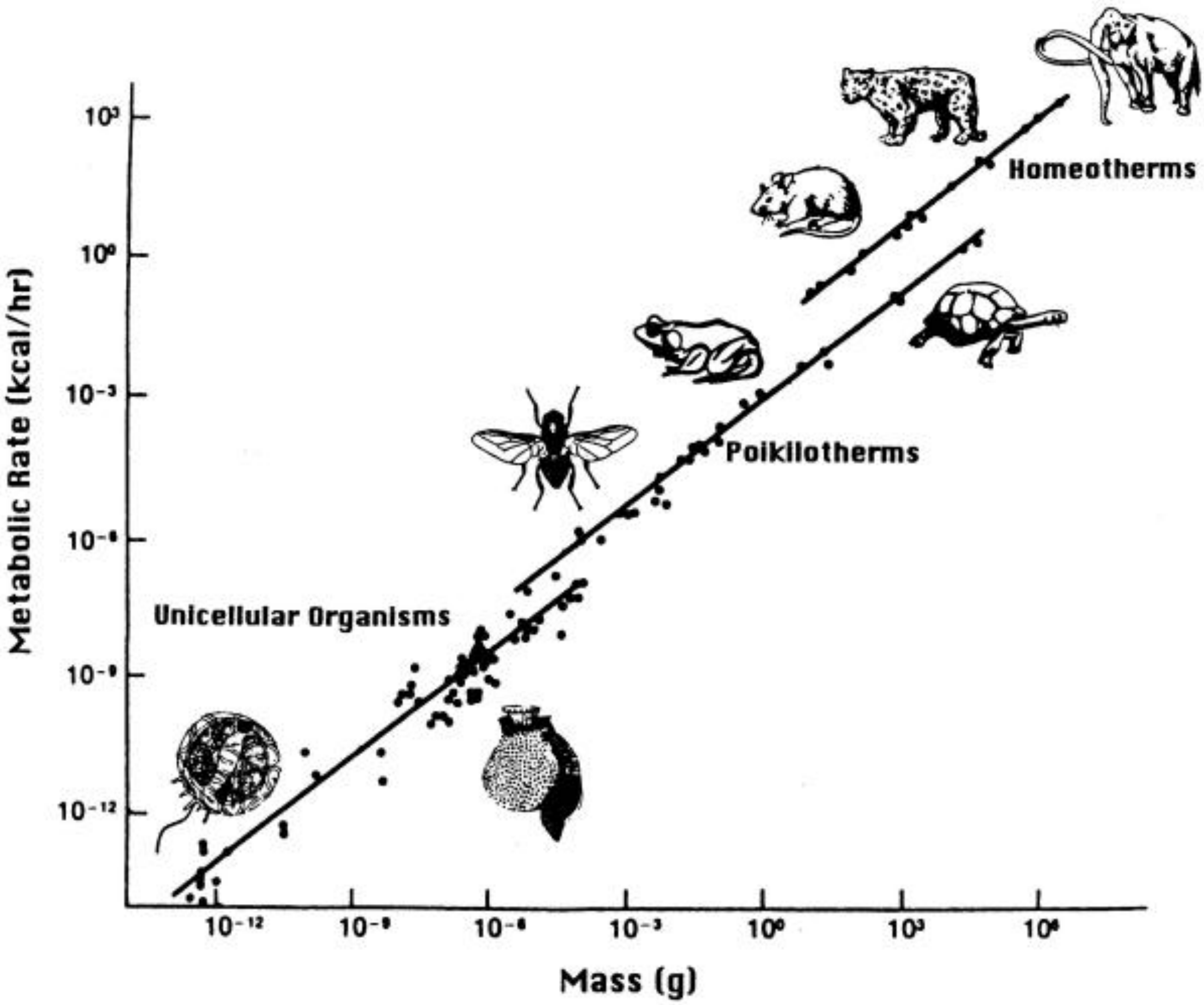


Figure 1: Relationship among Mass and observed Metabolic Rates[1]

Nature is scalable. Inspired by this phenomenon, a number of scalability studies have been conducted by the researchers all over the world to find the relationships among different parameters. One of the most prominent among them is Moore's law [3]. Moore's Law is a scalability study which established a relationship among the number of years and the growth in number of transistors per electronic chip. The law indicated that number of transistor in a dense electronic circuit will double approximately every two years as shown in Figure 2.

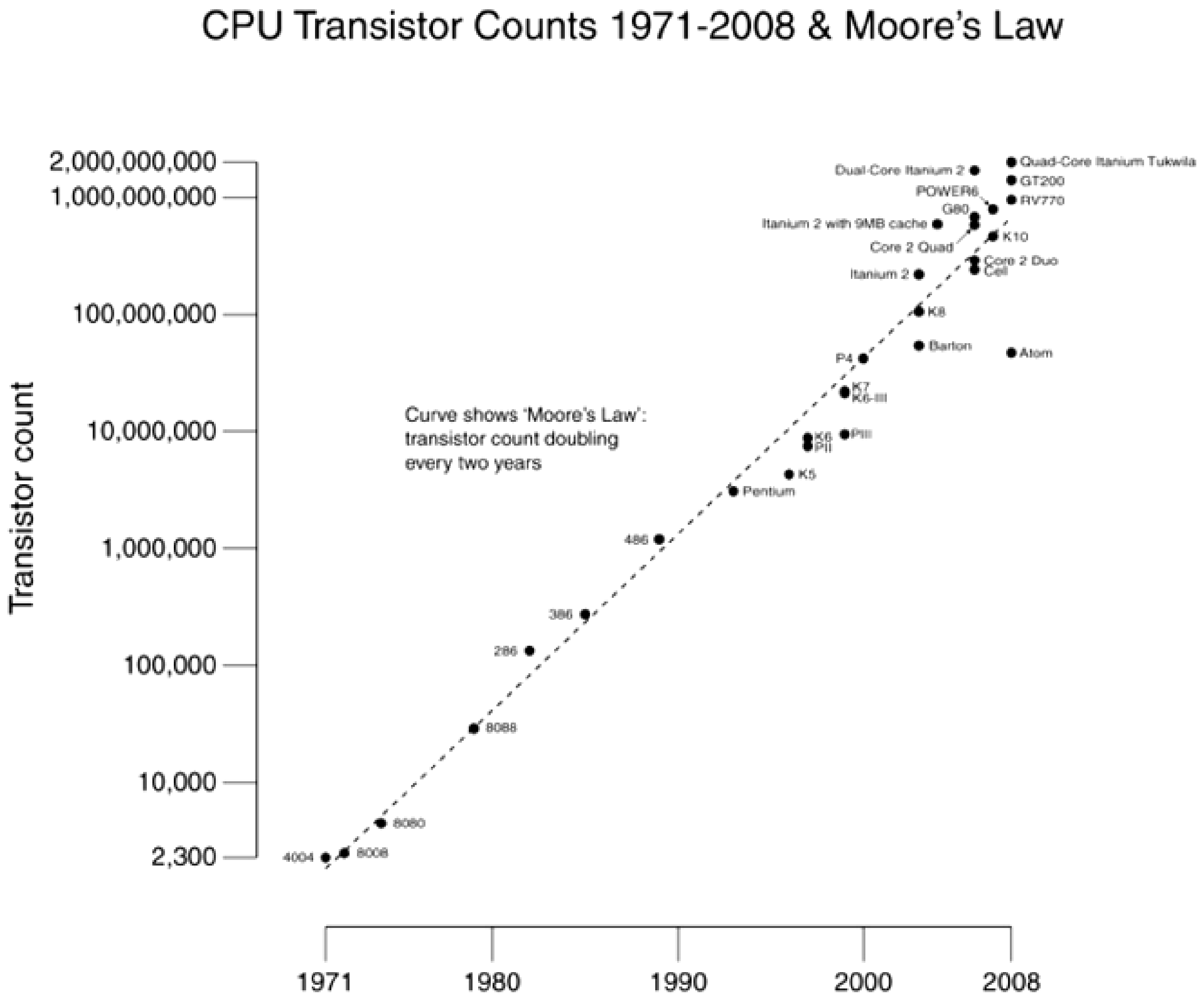


Figure 2: Moore's law deductions on increase of transistor counts [3]

Another study explains the basics of flight and establishes the scalability among different flying creatures and machines. In great flying diagram (Figure 3), we can see that smaller flying animal/machine have scalable relation with bigger animals and machines. The diagram is constructed using the weight and wing loading against cruising speeds. The deductions are unique and an eye opener as it opens the door for unique aviation designs which are yet to be absorbed by humans but have existence in nature. He further mentions that there are certain convergence and divergence to the scalability; however there diversity can be explained, but good designs are those that are close to the line. The designs that are away from the line need either more muscle strength or power to remain in air, thus they either have low endurance or are muscle exhaustive/maintenance intensive, His report further explains each flying animal's categories, in order to jog up the readers mind to look for nature inspired designs which will induce more confidence in future aviation designs [5].

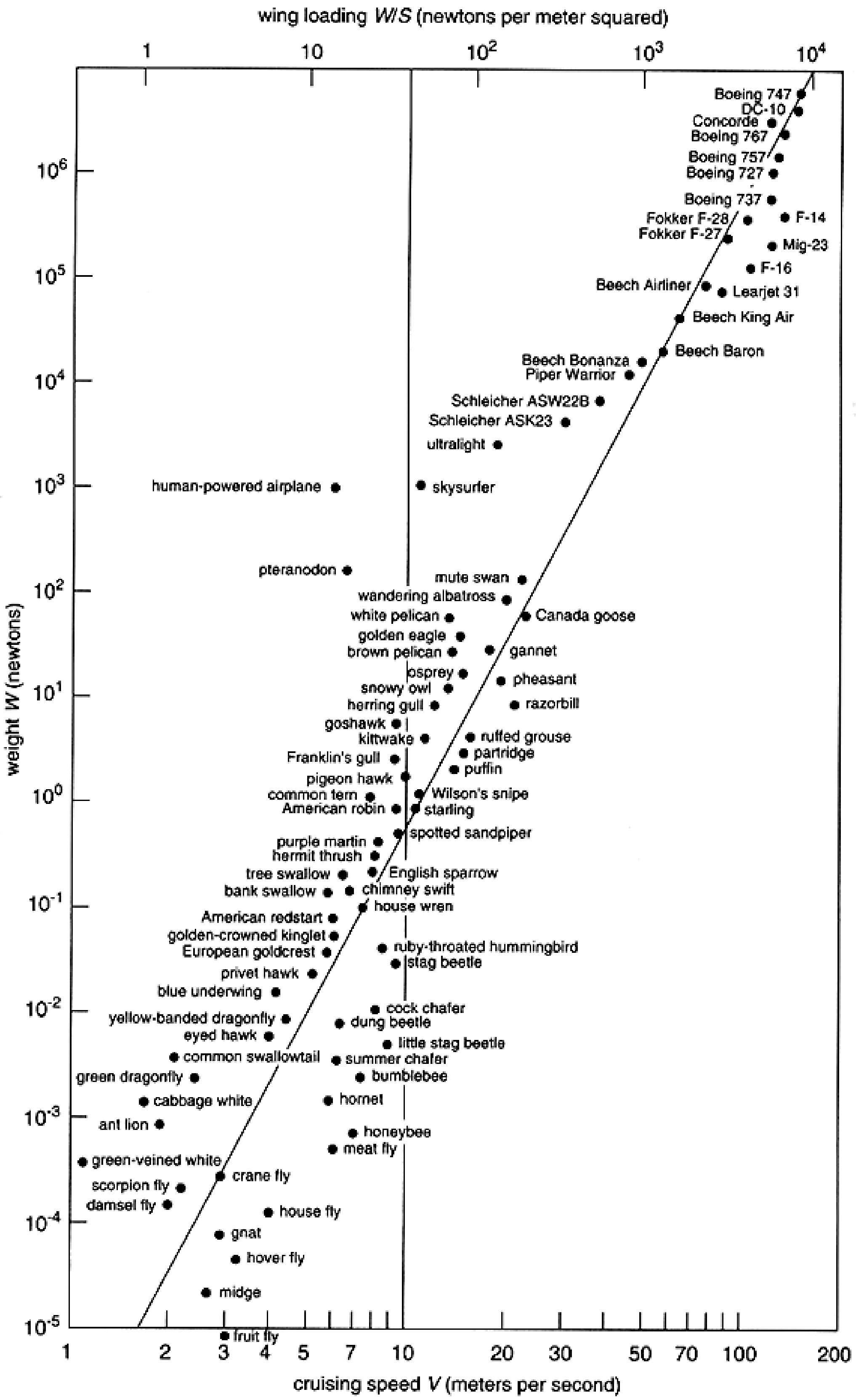


Figure 3: The Great flying diagram[5]

2.2 Surrogate modeling in complex systems

The above mentioned studies were all related to scalability relations among nature and technologies. The scalability relations are studied among two parameters. These scalability studies depict the interdependency of different parameters. However there are cases in which many parameters are predicting the response of certain variables. For such cases different studies are performed using regression techniques. The use of surrogate models in high fidelity systems has improved the overall performance of these systems. These models are also regarded as a system engineering tool to help funnel complex systems towards optimization [6]. In estimation studies, surrogate models have been developed for predicting floods [7], evaporation rate estimation in metrological field [8], analyzing customer satisfaction for airlines [9] , predicting airline delays [10] and indirect costs [11].

2.2.1 Application of surrogate modeling in Aviation Industry

A study was conducted at School of Aeronautics, Northwestern Polytechnical University China, where surrogate models were studied [12]. It's a detail study in which different surrogate techniques are discussed. Surrogate models are used to find global or local optimization. They intend to optimize the process with the use of models whereas the conventional process includes time exhaustive analysis codes. However it is to be noted that surrogate base optimization is only approximation to true optimum that is why these surrogate models need to be updated on regular interval with the inclusion of new sample points. A comparison of conventional and surrogate based optimization is mentioned in figure 4.

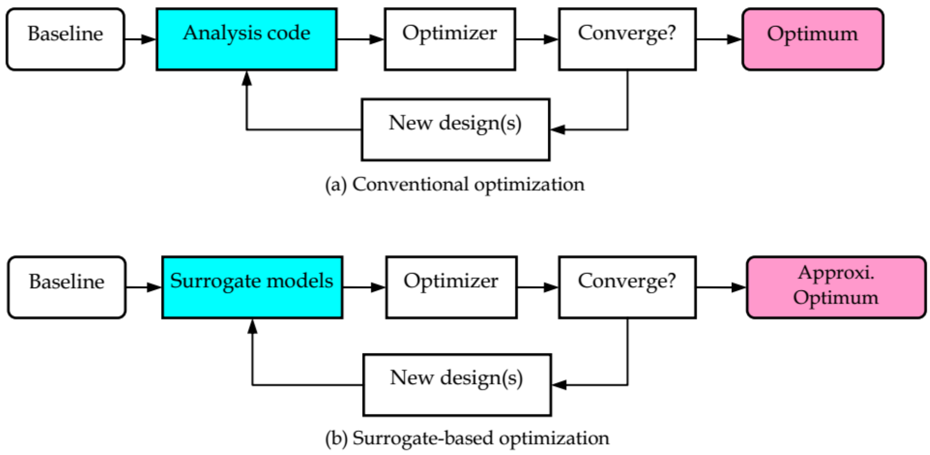


Figure 4: Comparison of frameworks of conventional optimization and surrogate based optimization [12]

The researcher has discussed response surface method (RSM), krigging and radial base function (RBF). RSM is a polynomial approximation model in which the sampled data is fitted by a least-square regression technique whereas krigging is an interpolation of observed data. RBF is an alternative interpolating technique whose value depends on the distance from the origin. These models are explained in detail along with their applicability. The author in section of application of these models has estimated airfoil and wing designs. Similar study was carried out in estimating aerodynamic coefficient and drag polar for airfoil RAE 2822. In this study, two krigging approaches were combined to improve the overall accuracy of existing Variable Fidelity Model (VFM) approach [13].

Another study on surrogate modeling has highlighted the use of surrogate modeling in estimating the performance of an inverted wing with counter-rotating vortex generators in ground effect. Technique involve to build such models is co-krigging which utilizes the experimental (wind tunnel data) and computational (computational fluid dynamic software's) data to interpolate vortex generation. In addition to that it was analyzed that how much a low fidelity data contributes in improving surrogate models than using limited high fidelity data. The results have shown that use of such multi fidelity data has improved the model performance [14].

The use of multiple linear regression approach has also been seen in an unclassified report by Aviation research laboratory of university of Illinois in 1978. The predictor display used for aircraft simulators was discussed. This predictor display system shows the human input to aircraft flight path and its results in future. The existing techniques to calculate the six degree of freedom of equations of motion in fast time are fairly inaccurate. The use of statistical approach to calculate predictor information has found out to be fast, low cost and reliable. The researcher presents least-squares, regression approach for determining first-order, and linear approximations of accurate fast-time models used in predictor displays. Such a procedure would eliminate the need for an operational fast-time model while still providing a great deal of predictive accuracy. However further research will be needed to able to implement it safely and effectively [15].

An interesting study was carried out using statistical methods to evaluate aircraft trajectory and then comparing it with existing air traffic management techniques. The research includes standard in use software results compared with estimation techniques. Multiple linear regression along with other statistical methods was applied over existing data set. On-board management systems predict the aircraft trajectory using a point mass model of the forces applied to the center of gravity. The point-mass model requires knowledge of the aircraft state (mass or thrust), atmospheric conditions (wind or temperature), and aircraft intent (target speed or climb rate). Many of this information is not known to the ground based systems and the available information is known with certain accuracy. Both linear and nonlinear frameworks of regression analysis are used in this research. The methods employed are multiple linear regressions along with principal component analysis (PCA), regression using neural networks (NN) and Loess method. These applied methods are then compared along with the existing BADA model results and are shown in figure 1. The Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) can be seen in above table against each method (smaller the value better fitted is the model). The values for regression techniques are significantly less than the existing BADA model [16].

Method	MAE	RMSE
BADA	1140(79)	1824(95)
Least Square	744(55)	962(72)
Neural network	841(47)	1080(55)
Loess Method	699(54)	908(72)

Table 1: Comparision of different models

Alan J Stolzer [4] conducted a study on calculation of fuel consumption of commercial aircraft using regression analysis. He emphasized on the fact that 10% of commercial aircraft operational budgets is consumed by fuel and if the fuel management is overlooked the percentage can be much higher and can contribute to financial losses. The fuel consumption reading taken at different check points during the flight path was used to build statistical models using regression analysis. This technique can help operators to plan their route with more efficiently and effectively.

Similar study was conducted at University of Toronto, where a multi-mission fuel-burn minimization was performed. The aerodynamic and aero structure design optimization was studied for different flight paths and missions. It is said that a design needs to be optimized for a set of flight profiles only then that design is an effective design. A long-range aircraft configuration similar to the Boeing 777-200ER with the objective of minimizing the weighted average fuel burn of a large set of missions was selected. The resulting aero structural optimization had 311 design variables and 152 constraints, and it solved for a total of 28 flight conditions at each optimization iteration[17].

Surrogate models have also been developed to aid decision making processes. The models have been applied to predict aircraft emissions that are being developed to support aviation environment policy making. The interesting deduction were how confidence intervals can be helpful in tradeoff between computation time and uncertainty in the estimation of the statistical outputs of interest[18]. Similar research has been carried for a case study of space propulsion: a response surface-based multi-objective optimization of a radial turbine for an expander cycle-type liquid rocket engine. The model is combined with a genetic based algorithm and is found effective in supporting global evaluations. The study however lacked the established experiences in adopting radial turbines for space propulsion, so designer intuitive ideas were employed to build surrogate models for conduct of optimization framework which helped in construction of genetic algorithms [19].

The latest research of use of scalability relation and regression is seen by Adnan Ashraf, a graduate student of National University of Sciences and Technology (NUST) in 2016. His research was centered on Unmanned Aerial Vehicles (UAV), in which he presented scalability relations and certain regression models. A detail study was carried out of his research before starting the current research and certain shortcomings were noted in his research and are mentioned below [20].

i. The data collected was missing certain parameters that are necessary for estimation such as element of power / thrust was missing during model building.
ii. The researcher could not establish the link between power law models and regression models, as to which models should be selected and what are the merits of selection.
iii. The regression and scalability tools can only be applied in line with the subject knowledge; we cannot predict a variable using those predictor variables which are contrary to the fundamentals of aviation.

These issues have been duly catered during this research

2.3 Existing analytical equations for aircraft performance estimation

Surrogate models are developed for aircraft performance parameters that are maximum velocity, range, service ceiling, rate of climb and endurance. There brief description along with analytical equations are available in Anderson [21]. The surrogate based study has been classified into two broad categories of turbojets/turbofan and propeller driven aircraft. The propeller driven aircraft also include battery driven propeller aircrafts.

2.3.1 Maximum Velocity (V_{max})

It's the maximum attainable velocity by an aircraft under certain conditions (such as altitude and propulsive property).

V_{max} for a jet aircraft in steady straight flight can be determined using

$$V_{\max} = \left[\frac{[(T_A)_{\max}/W](W/S) + (W/S)\sqrt{[(T_A)_{\max}/W]^2 - 4C_{D0}K}}{\rho_\infty C_{D0}} \right]^{1/2} \quad (1)$$

where T_A is thrust available, W is total weight, S is wing surface area, ρ s density, C_{D0} is coefficient of drag at zero lift and k is constant for coefficient of drag due to lift. In above equation *$(T_A)_{max}/W$* increases, *W/S* increases and *C_{D0}* and/ or *K* decreases

For propeller aircraft the maximum velocity is determined graphically using general equation of

$$P_A = T_A V_\infty \quad (2)$$

2.3.2 Rate of Climb (RoC)

In aviation, rate of climb (RoC) is vertical speed – the rate of positive altitude change with respect to time or distance. The analytical equations are

For Jet propelled aircraft

$$(R/C)_{\max} = \left[\frac{(W/S)Z}{3\rho_\infty C_{D0}}\right]^{1/2} \left(\frac{T}{W}\right)^{3/2} \left[1 - \frac{Z}{6} - \frac{3}{2(T/W)^2 (L/D)_{\max}^2 Z}\right] \tag{3}$$

where

$$Z = 1 + \sqrt{1 + \frac{3}{(L/D)_{\max}^2 (T/W)^2}} \tag{4}$$

where T is thrust of aircraft, W is total weight of aircraft, S is wing surface area, $\mathscr{P}$ is density, C_{D0} is coefficient of drag at zero lift and k is constant for coefficient of drag due to lift.

For Propeller driven aircraft

$$(R/C)_{\max} = \frac{P_A}{W} - \frac{2}{\rho_\infty}\sqrt{\frac{K}{3C_{D0}}}\left(\frac{W}{S}\right)^{1/2} \frac{1.155}{(L/D)_{\max}} \tag{5}$$

where P_A is power available, W is total weight of aircraft, S is wing surface area, ρ is density, C_{D0} is coefficient of drag at zero lift and $(L/D)_{max}$ is the maximum lift to drag ratio.

2.3.3 Service Ceiling

The maximum height at which a particular aircraft rate of climb drops to zero is called absolute ceiling whereas usable height at which the aircraft can sustain a steady rate of climb is called service ceiling (this is normally at *(R/C)*$_{max}$*=100 ft/min*).

Procedure to obtain ceilings is to first calculate *(R/C)*$_{max}$ at various different altitudes then plot *height* vs *(R/C)*$_{max}$ and finally extrapolate the curve to *(R/C)*$_{max}$ *=0* and *(R/C)*$_{max}$*=100ft/min* to get the absolute and service ceilings respectively.

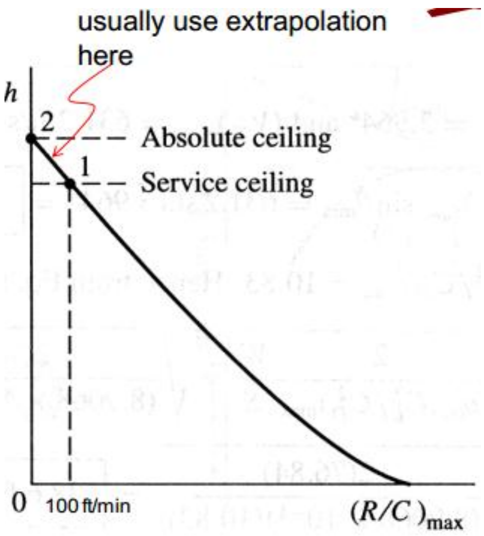


Figure 5: Absolute and service ceilings

2.3.4 Range

It is the maximum distance an aircraft can fly between takeoff and landing.

For Jet propelled aircraft

$$R = \frac{V_\infty}{c_t} \frac{L}{D} \ln \frac{W_0}{W_1} \qquad (6)$$

where V_∞ is free stream velocity, L/D is lift to drag ratio, W_0 is gross weight of aircraft and W_1 is the aircraft without fuel weight and c_t is thrust specific fuel consumption. In above equation range is influenced by fuel capacity, L/D, V_∞ and engine efficiency or thrust specific fuel consumption (TSFC) c_t , higher range can be achieved by higher fuel capacity (W_0/W_1). Higher L/D , higher V_∞, and lower c_t, V_∞ and L/D are not independent and to maximize range, aircraft needs to fly at condition where V_∞ (L/D) is at maximum.

For propeller driven aircraft

$$R = \frac{\eta_{pr}}{c} \frac{L}{D} \ln \frac{W_o}{W_1} \qquad (7)$$

In above equation aircraft needs to fly at maximum L/D with maximum propeller efficiency (η_{pr}), minimize Specific fuel consumption(SFC) c and maximum fuel capacity (maximize W_0/W_1) in order to get maximum range.

The expression for electric powered propeller planes is different from the engine driven propeller, as the Breguet equation doesn't cater for the battery behavior and its effective capacity (Peukert effect). The equation for range for a battery driven propeller is

$$R_{\max} = Rt^{1-n}\left(\frac{\eta_{total} VxC}{(1/\sqrt{\rho S})C_{D0}^{1/4}(2W\sqrt{k})^{3/2}}\right)^{n} \sqrt{\frac{2W}{\rho S}\sqrt{\frac{k}{C_{D0}}}}\, x3.6 \qquad (8)$$

where Rt is battery hour rating in hours, η_{total} is total efficiency, V is voltage , C is capacitance of battery, S is surface area of wing , ρ is density, W is total weight , C_{D0} is coefficient of drag at zero lift and k is constant for coefficient of drag due to lift.

2.3.5 Endurance

It's the maximum time the aircraft can spend in cruising flight. Endurance is also affected by engine performance through specific fuel consumption (SFC) or thrust specific fuel consumption (TSFC).

For Jet propelled aircraft

$$E = \frac{1}{C_t}\frac{L}{D}\ln\frac{W_0}{W_1} \qquad (9)$$

To attain maximum endurance aircraft needs to fly at maximum L/D with minimum TSFC (c_t) with maximum fuel capacity (maximize W_0/W_1)

For Propeller driven aircraft

$$E = \frac{\eta_{pr}}{c}\sqrt{2\rho_\infty S}\,\frac{C_L^{3/2}}{C_D}(W_1^{-1/2} - W_0^{-1/2}) \qquad (10)$$

To achieve maximum endurance aircraft needs to fly at maximum $C_L^{3/2}/C_D$, fly at sea level (maximum ρ_∞) with maximum propeller efficiency (η_{pr}), minimize SFC (c) and maximize fuel capacity (maximize ($W_1^{-1/2}$-$W_0^{-1/2}$)).

The endurance calculation for a battery driven propeller aircraft is different than.. (what) .

The equation of endurance is then transformed into

$$E = Rt^{1-n}\left(\frac{\eta_{total}VxC}{(2/\sqrt{\rho S})C_{D0}^{1/4}(2W\sqrt{k/3})^{3/2}}\right)^n h \qquad (11)$$

where Rt is battery hour rating in hours, η_{total} is total efficiency, V is voltage , C is capacitance of battery, S is surface area of wing , ρ is density, W is total weight , C_{D0} is coefficient of drag at zero lift and is constant for coefficient of drag due to lift.

2.3.6 Analytical equations calculation - An intricate process

A detailed overview of the aircraft performance parameters are presented in above equation from (1) to (11) along with different propulsion systems. Following are some of the observations on this analytical approach:

a) Each propulsion system needs to be catered separately during analytical calculations. The equations are different for turbojets, propeller and battery powered aircrafts.

b) As evident from each equation, one needs detailed data of aircraft which include data from geometric, propulsion (battery specification for battery driven aircrafts), wing design, wind tunnel data for wing / aircraft and atmospheric properties.

c) Data needed for performance parameters calculation is calculated via time and cost exhaustive process such as wind tunnel testing or propulsive properties including specific fuel consumption or thrust / power available calculations.

d) Moreover these analytical equations are still approximations as these equations are simplified using many assumptions such as neglecting different environmental effects or other thermal and structural phenomenon. Even wind tunnel test are not as what aircraft experience in real world environment. These assumptions are necessary so as to simplify the initial calculation process along with design refinement / optimization at each step.

2.4 Missing link in Literature

Statistical studies have been conducted in almost every complex field of life and technology and it yielded significant results. Surrogate models developed using statistical techniques are also applied in aviation field. However the application use has been limited to maintenance and operational aspect of aviation industry as explained in above literature review. The use of surrogate modeling to benefit the designer group has been very limited; in fact the recent research by graduate of National University of Sciences and Technology (NUST), had its own gray areas that needed to be addressed in order to build fitted surrogate models to ease the designer current methodologies. The gray areas of his study have already been mentioned above.

Moreover till date no comprehensive study has been conducted to give conclusive findings which can be helpful for design engineers. We have studied use of surrogate modeling in different aviation fields but no contributions have been made in estimating aircraft performance parameters. If aircraft performance parameters can be estimated using validated surrogate models, it will help the designers to create initial design bounds for their projects.

CHAPTER 3

METHODOLOGY AND PROBLEM FORMULATION

3.1 Introduction

The designers over decades have come up with different aircraft designs following some fundamental principles. We have now sufficient amount of data for different aircraft for which we can generate approximate / surrogate models from earlier designs. These surrogate models can act as predictors of futuristic design. Although the results of using such models might be crude but the time saving outstrip this limitation. Moreover, these models can help novice designers to generate different estimates without going into intricate details. It is proposed that if statistics based simple power laws and regression techniques are developed based on primitive information, several performance parameters can be approximated with higher confidence. This approach will help to circumnavigate the tedious processes of detailed geometric modeling and extensive involvement of aerodynamic and propulsion information. The comparison of traditional and proposed approach is depicted in Figure 6. The advantage of using proposed approach is evident from Figure 6. The detailed geometric modeling, followed by aerodynamic, propulsive and flight performance equation formation can be replaced by using statistical approach and estimate flight performance parameters in an efficient and effective manner.

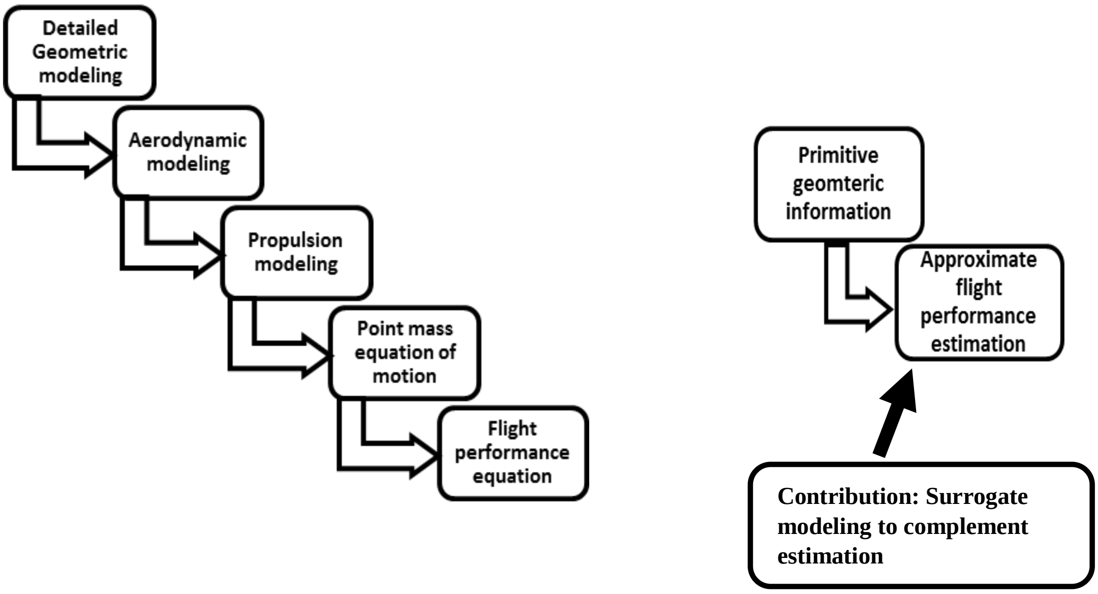


Figure 6: Conventional design methodology vs Surrogate model based methodology

The adopted methodology for the conduct of this thesis is shown in Figure 7, where the process is initiated with the need analysis phase. After defining the need which is to build surrogate models for aircraft performance parameters approximation, a data set was collected upon which further statistical tools will be applied. The models developed will then be validated using different criteria's and finally the validated model will be called surrogate model for aircraft performance estimation. The software utilized in models building is Matlab® for single variable models using power laws and Minitab® V16 for regression analysis in order to form multiple variable models. There are different techniques to build surrogate models, however after careful scrutiny and studying the parameters behaviors, scalable relationship were developed using power laws and multiple linear regression (MLR). The use of these techniques in high fidelity engineering solution is highly effective as it helps to reduce the design cycle time and cost by enabling rapid analysis of alternative designs.

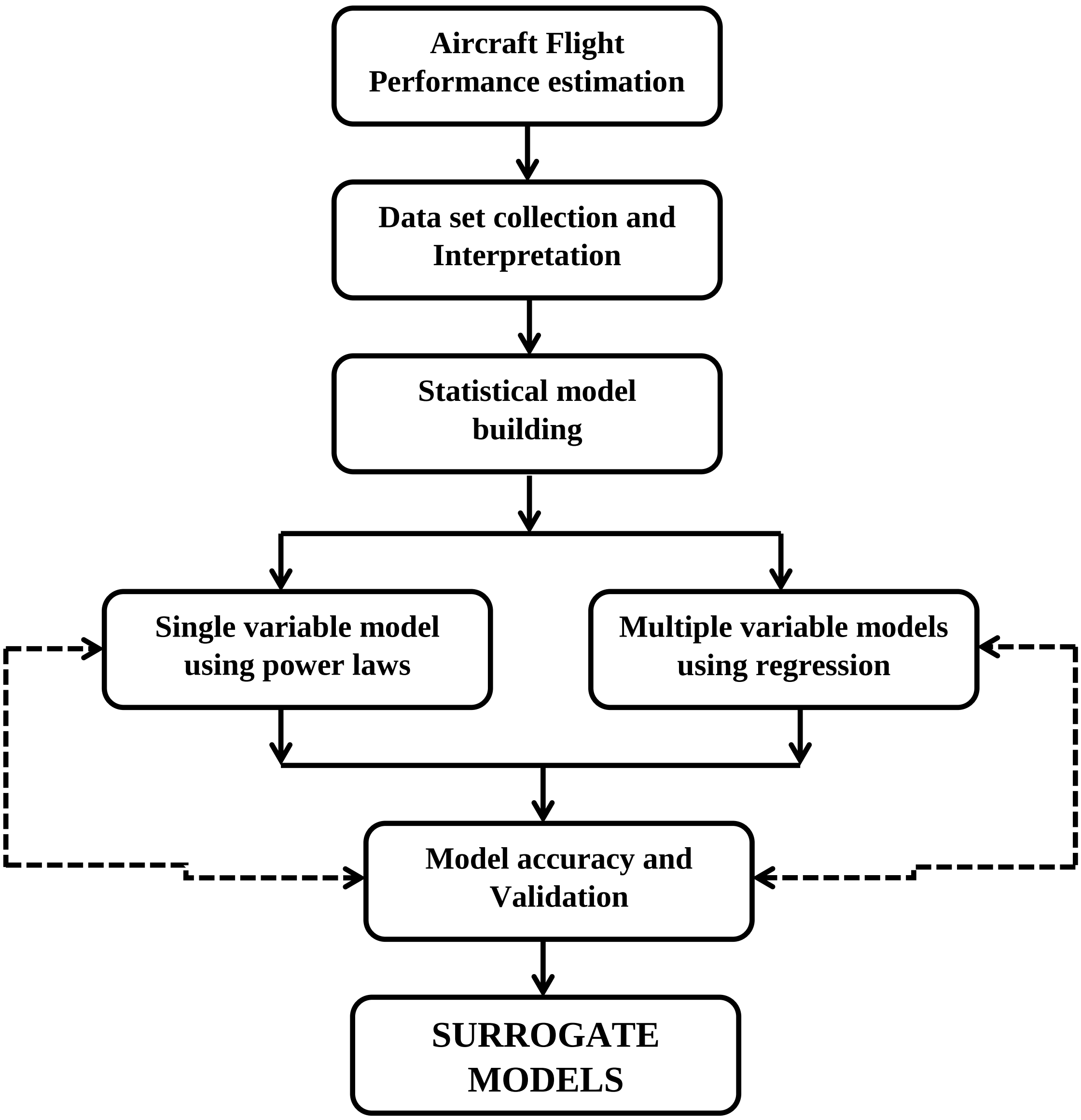


Figure 7: Methodology of research

3.2 Data collection and interpretation

For the analysis purposes a data collection study was conducted to collect different types of aircrafts ranging from different eras. During data collection, **Janes All the world aircraft 2007-2008** volumes were very helpful along with other aircraft encyclopedias [22-24]. The following table gives an overall summary of the data collected. The aircrafts were broadly categorized as unmanned air vehicle (UAV), jet fighter and airliner aircrafts. Further for UAV the data set includes all types as high altitude long endurance (HALE), medium altitude long endurance (MALE), tactical UAV (TUAV), mini aerial vehicle (MUAV), micro aerial vehicle (MAV) and unmanned combat air vehicle(UCAV). On the same lines Jet fighter aircrafts is segregated into generations which depend on aircraft design, avionics, and weapon system. It is to note that generation shift in jet fighter aircraft occurs when a technological revolution cannot be incorporated into an existing aircraft through upgrades or retrofits. In the airliners category all in-service airliners were included which encompasses different designs. The overall data set summary is in Table 2.

No.	Aircraft Category	Count	Remarks
1	Unmanned air vehicle	148	Includes MALE,HALE,MAV & UCAVs
2	Fighter Aircrafts	82	Includes all generations of aircrafts
3	Airliners	67	Includes all current used airliners
Total		297	

Table 2: Summary of data collected

For each aircraft certain parameters were collected, the collected parameters have been classified into geometric/design parameters and performance parameters. The consistency of units was carefully maintained while data collection. The estimated models can only be developed with high confidence if the units are consistent across the data set. The parameters collected are shown in Table 3. The idea is to predict aircraft performance parameters using geometric parameters. The geometric parameters include basic dimensional parameters and other basic design parameters, these parameters are easily available over internet sources as mentioned. Aircraft performance parameters do require certain parameters for calculation, which requires exhaustive experimental analysis and wind tunnel testing. As mentioned earlier, if the calculations are circumvented using basic parameters, it will open new horizons for design engineers. It is to be noted that for certain parameters were collected only for certain type of aircraft due to non-availability of information by OEM. Endurance data was collected only for propeller based UAV's and Rate of climb (ROC) data for propeller based UAV's was fewer.

	Parameters Collected	
No.	**Geometric / Design Variables**	**Performance variables**
1	Length	Maximum velocity
2	Wingspan	Range
3	Wing Area	Ceiling
4	MTOW	Endurance
5	Empty Weight	Rate of climb
6	Thrust/Power	

Table 3: Parameters collected

3.3 Single variable model using Power laws

The power law is used to find one to one relationship among parameters. Using "Y" as dependent variable and "X" as independent variable, the power law scaling is

$$Y = \alpha * X^{\beta} \qquad (12)$$

where α is normalization constant and β is power of independent variable. The power law equation gives two important insights that need to be kept in mind while exploring the relation among parameters.

i. The dependent variable ‘Y’ is normalization constant ‘α’ times the dependent variable ‘X’ with raise to the power of ‘β’

ii. Secondly and very important deduction is the dependence of β value for estimation of dependent variable (β>1, it means the relation is super linear, β=1 means linear and β<1 means sub linear). This means β value will determine whether the dependent variable increase radically or normally.

3.4 Multiple variable models using Regression technique

Regression Analysis is a statistical process for estimating the relationship among variables. It includes many techniques for modeling and analyzing several variables, when focus is on the relationship between a dependent variable and one or more independent variables. Regression Analysis helps in understanding how dependent variable changes when one of the independent variable is varied. The simplest regression model is the simple linear regression model, which is written as:

$$y_i = \beta_0 + \beta_j x_i + \varepsilon_i \tag{13}$$

where “y” is dependent (response) variable, β_0 is intercept (mean of dependent variable when x is zero), β_j is slope (change in **y** w.r.t **x**) or regression coefficients and ε_i (Random part) explains variability of response about the mean. This regression analysis is a set of procedures based on a sample of “n” order pairs (x_i, y_i), i= 1, 2, 3,…., n, for estimating and making inferences on the regression coefficients, j=1,2,3,….,n for multiple independent variables regression coefficient. These estimates can then be used to estimate mean values of dependent variables for specified value of x. Various diagnostics checks have been used to assess the quality of the developed model(s) and its resultant estimates such as Anderson Darling test, R-squared, Adjusted R-Squared, t-test and f-test.

3.3.1 R-Squared

R^2 designates the proportion of variance in the dependent variable which is predictable from independent variable. It is a statistical test, used in the perspective of statistical models with the purpose of either the prediction of future consequences or the testing of hypotheses based on the related information. It provides the measure of accuracy of observed outcomes as replicated by the model. The R^2 value ranges from 0 to 1. If the data set have n values identified as [9] each associated with a predicted value $\{f_1 < \ldots < f_n\}$. The R^2 value is calculated by

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \qquad (14)$$

where sum of Square of residuals (SS_{res}) is given by

$$SS_{res} = \sum_i (y_i - f_i)^2 \qquad (15)$$

And the total Sum of Squares (SS_{tot}) by

$$SS_{tot} = \sum_i (y_i - \bar{y})^2 \qquad (16)$$

3.3.2 Adjusted R-Squared

To measure how successful the fit is in terms of explaining variation of data. Adjusted R^2 is just a change of R^2 that adjusts the amount of terms in a statistical model. Adjusted R^2 calculates the proportion of the variation in the dependent variable caused by the predicting variables. The adjusted R^2 is calculated by

$$R^2_{Adjusted} = 1 - \frac{(1-R^2)(N-1)}{N-p-1} \qquad (17)$$

where "R^2" is the sample R squared value, "p" is number of predictors and "N" is total sample size.

3.3.3 t-test

t-test is a statistical hypothesis test, in which test statistics follow a t distribution under the null hypothesis. It is used to conclude if two sets of given data are considerably different from each other. It is most frequently applied when the test statistics follow normal distribution and the value of scaling term is known. If the scaling term is not known then its substituted by an estimate based on the given data. So it is used to assess significance of the individual regression coefficients. The t value is calculated by

$$t = \frac{M_x - M_y}{\sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}} \qquad (18)$$

$$S^2 = \frac{\sum (x - M)^2}{n - 1} \qquad (19)$$

where "*M*" is mean, "*n*" is number of score per group and "*x*" is individual score.

3.3.4 F-test

It is a statistical test where the test statistics have an F-distribution under the null hypothesis. It is most frequently used is the comparison of statistical models, fitted to given data set, for identifying the model that best fits the population from which the data is sampled. Precise f-tests results when the models have been fitted to the data using least squares. It is used to assess the overall adequacy of the model. F value is stated as the ratio of variances of two observations. The association between the variance of two data sets can lead to many estimates. The formula for F test is

$$F_{value} = \frac{\sigma_1^2}{\sigma_2^2} \qquad (20)$$

where σ^2 is the variance and given by

$$\sigma^2 = \frac{\sum (x - \bar{x})^2}{n-1} \qquad (21)$$

where "x" is the given value, "x¯" is the mean value and "n" is total number of terms.

CHAPTER 4

SINGLE VARIABLE SURROGATE MODELING

4.1 Background

After the data collection process, the collected parameters behavior needed to be studied. Individual parameters behavior among each other will give us the insight as to how the models can be developed. In this preview, single variable models were developed so as to find out scalability trends among aircraft parameters.

4.2 Scalability relations

All aircraft look alike in terms of wings and fuselage and other notable features but they all vary in sizes. Now if we compare two aircrafts with one having twice the weight of other, then we will notice that it is not only the weight that has increased, the aircraft needs bigger wing size and other geometric alterations, to be able to make it fly worthy. A scalable relation among the two aircraft can be deduced that if the weight is increased by one unit then correspondingly wing loading and other geometric / performance parameters needs alteration to be able to fly and performs the given set of tasks. Scalable relations were studied among above stated parameters. This scalability study will form the foundation for further building of surrogate models of aircraft performance parameters as it will give a clear picture of one to one behavior among parameters.

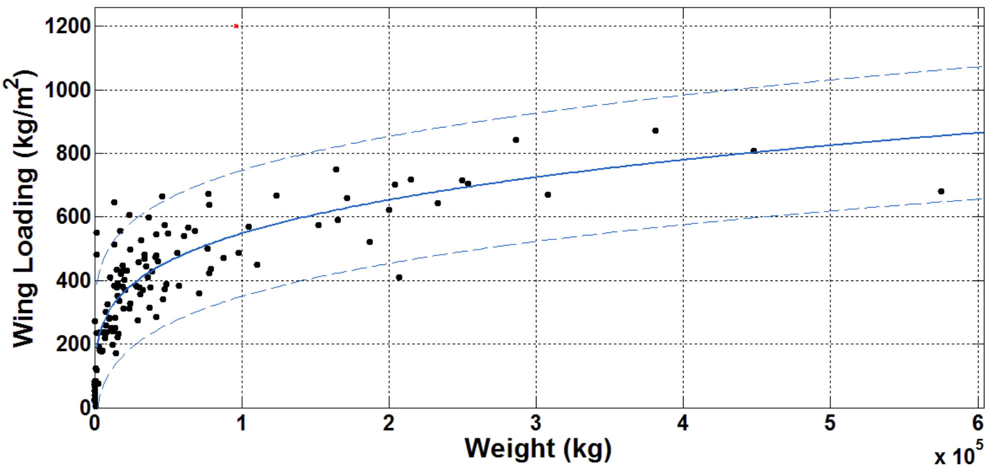
In order to form scalable relations, the one to one graphs were plotted among different parameters. A curvilinear trend was observed among parameters as shown in Figure 2 graphs. To better study such curvilinear pattern, different equations (Fourier, polynomial, exponential, Gaussian and power law) were implemented on a given set of data. It was found out that only power law equation best fits the aircraft geometric and performance data set.

The power law is used to find one to one relationship among parameters. Using “Y” as dependent variable and “X” as independent variable, the power law scaling is

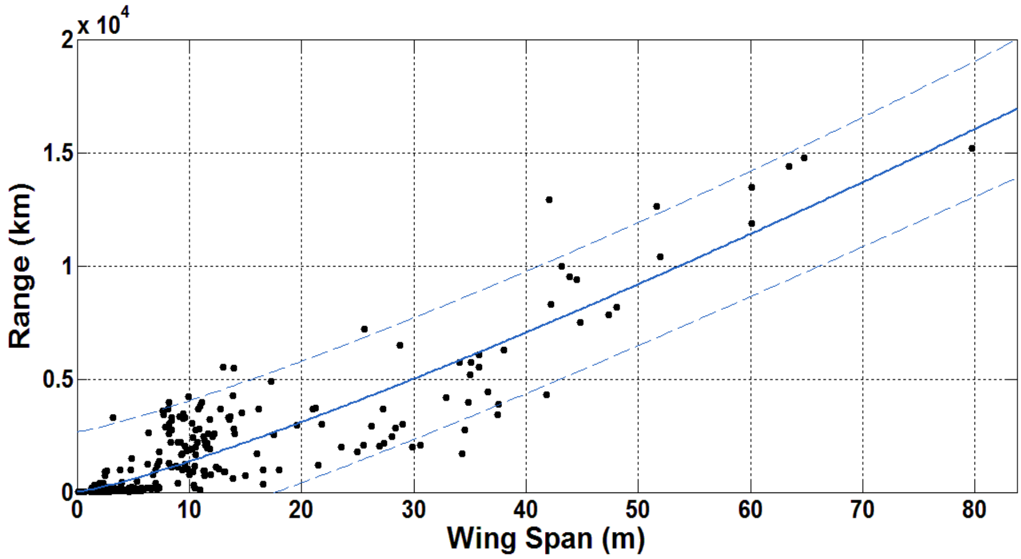
$$Y = \alpha \cdot X^{\beta} \tag{22}$$

MATLAB® curve fitting tool was used to perform power law analysis. The obtained results reveal that scaling trends exist among aircraft variables. Robust and proportionate scaling exponents have been found for different geometric and performance parameters. If one unit of predictor (X) is increased then alpha units times value will be the Response (Y). This gives us a very unique insight into the scalable relationships of aircraft parameters. For example if we take the case of weight of aircraft versus the wing loading (W/S), by applying the power law on weight vs wing loading data set, the scalable relation is $W/S = 29.35 * W^{0.25}$. This means that if weight is 1 unit then wing loading is 29.35 time 1 unit to power units. Similar scalable relations have been found among different parameters and are shown in the following table. The 95% confidence bounds give the tolerance for value. Notice the value for, super linear exponents will have an increase effect on dependent variable as compared to linear and sub linear exponents. This means that for cases such as range vs wing span , power vs weight and wing area vs wing span the range , power and wing area has a super linear relation with wing span, weight and wing span respectively however rest relationships are linear or sublinear in Table 4. Another interesting fact is that super linear relationship tend to have greater adjusted R^2 which means that the data set is best fitted for super linear cases, this is can also be seen in cases mentioned in Table 4. Further if the analytical equation study mentioned in Chapter 2 is compared with Table 4, interesting analogies are developed as the scalable relations are in line with the analytical equations. Range has a strong scalable relation with weight this is also evident from equation (6), (7) and (8). Similarly for other performance parameters, scalable relations can be compared with analytical equations for comparison. Power is dependent on weight of the aircraft, which means more weight requires more power to lift it. Likewise the heavier the aircraft , wing span and wing area are consequently greater in order to make it flight worthy and perform the intended tasks.

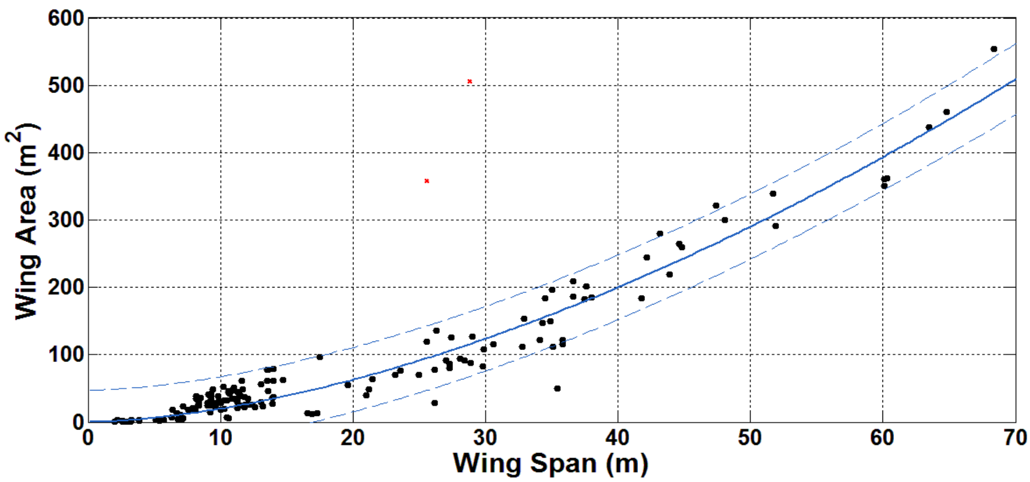
The scalability graph for Wing loading vs Weight is shown in Figure 8(a), Range vs Wing span in Figure 8(b) and Wing area vs Wing span in Figure 8(c). These all figures depict the existence of scalability among different aircraft parameters. The figures are labeled with 95% confidence bound, and maximum values are found in this region. The details of these graphs and other scalability relations are mentioned in Table 4, in respect of equation (1). The adjusted R^2 mentioned in Table 3, is a useful tool for comparing the explanatory power of models with different numbers of predictors. The adjusted R^2 will increase only if the new term improves the model more than would be expected by chance. The value for each explains the contribution of predictor for response variable (Super linear, linear and sub linear).



(a) Wing loading vs Weight with 95% confidence bounds



(b) Range vs Wing span with 95% confidence bounds



(c) Wing area vs Wing span with 95% confidence bounds

Figure 8: Scalable graphs of Aircraft parameters

Response (Y)	Predictor(X)	β	β , 95% Cl	Adj R^2	normalization constant, α	Obs
Endurance	Wing Span	0.8423	(0.7673, 0.9173)	78.58	1.9	117
Wing Span	Weight	0.451	(0.4257, 0.4762)	87.96	0.1948	278
Wing Area	Weight	0.8159	(0.7859, 0.8459)	96.14	0.014	161
Thrust	Wing Area	0.9093	(0.8485, 0.9702)	87.92	2.873	132
Range	Wing Area	0.6617	(0.5988, 0.7246)	76	193.3	155
Wing Loading	Weight	0.2542	(0.2241, 0.2844)	77	29.35	133
Ceiling	MTOW	0.1331	(0.1139, 0.1524)	55.6	3380	241
Range	Wing Span	1.186	(1.09, 1.282)	78.14	88.8	234
Range	Wing Area	0.6618	(0.5989, 0.7247)	76	193.2	155
Range	MTOW	0.5686	(0.5331, 0.604)	86.56	8.791	234
Power	Weight	1.076	(1.007, 1.144)	92.13	0.051	76
Thrust	Weight	0.822	(0.7834, 0.8607)	93.63	0.0249	172
Wing Area	Wing Span	1.669	(1.583, 1.755)	94.15	0.423	163
Range	Thrust	0.682	(0.629, 0.735)	82.28	115.9	154

Table 4: Scalability relationship of Aircraft geometric and Performance parameters

4.3 Concluding Remarks

These scalable relations are all logical with the basics of flight, these relationships also satisfy the analytical equations mentions in Chapter 2 section 2.3. Single variable models show a high value for adjusted R^2 and strong scalability trends have found among aircraft parameters. However the aircraft design process are complex and other modeling techniques can also be looked into for modeling of aircraft performance parameters. But these single variable models can be very much helpful in low fidelity design bounds.

CHAPTER 5

MULTIPLE VARIABLE SURROGATE MODELING

5.1 Background

As discussed in the previous chapter that single variable models although are fairly accurate, but aircraft design process require more than one variable model to depend upon due to the inherent complexity of design process. Surrogate models should include all such variables that can affect the design process. In this regard the multiple variable model techniques need to be looked into.

5.2 Multiple linear regression analysis

As mentioned earlier that power law explains the scalability among two parameters. However aircraft design is an intricate process, thus estimation via one parameter doesn't appeal to be much conclusive when initial estimation is a goal. If a parameter is estimated using more than one independent variable, the result will be much more conclusive and will help in significant cutting down of design iterations during initial stages. For example, if we need to find range of an aircraft and we use the wing span relation from above table, the estimates might be close for many inputs but it must be noted that if two aircrafts are having the same wing span, they will not have same range necessarily as range depends on other factors too. So the above scalable relations will only determine the design bounds but to better estimate the dependent variable, we need to find out relations that involve multiple independent variables. The multiple variable modeling requires further investigation, among other techniques such as multiple linear regression (MLR), krigging and radial base functions; MLR has been chosen as after data transformation (further explain in chapter 4 section), data depicted linear behavior. Krigging and radial base functions are generally employed once we encounter highly nonlinear design problems [8]. For this we employ multiple regression technique.

Many possible regression lines could be fitted to the sample data, but we choose that particular line which best fits that data. The best regression line is obtained by estimating the regression parameters by the most commonly used method of least squares. The principle of least squares consists of determining the values of the unknown parameters that will minimize the sum of squares of errors where errors are defined as the differences between observed values and the corresponding values predicted or estimated by the fitted model equation. The models are developed using the backward elimination method which is a process that starts with all predictors in the model and then eliminating the least significant variable in each step till the time p-value is greater than the specified limit. A pictorial illustration for said approach, use to build surrogate models is shown in Figure 10.

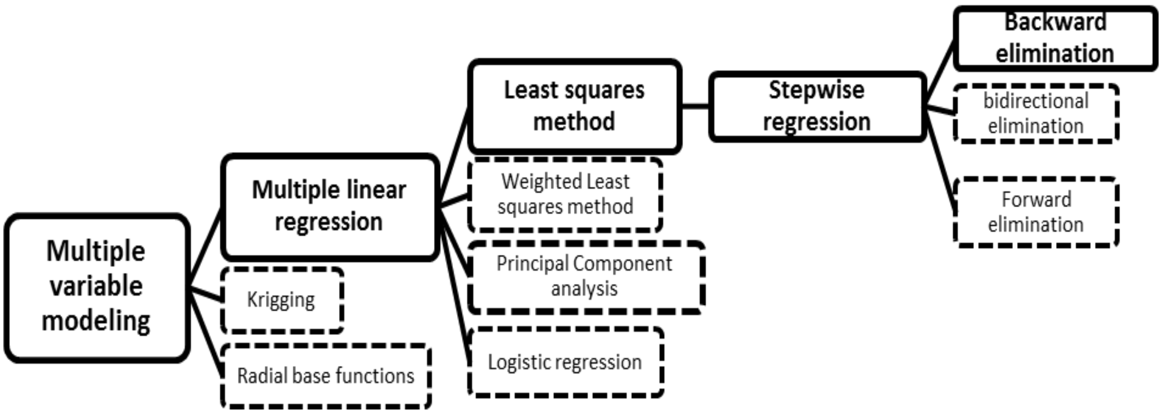


Figure 9: Multiple variable modeling methodology

5.3 Data transformation

First of all the data was transformed. In data analysis, transformation is the replacement of a variable by a function of that variable: for example, replacing a variable **x** by the square root of **x** or the logarithm of **x**. In a stronger sense, a transformation is a replacement that changes the shape of a distribution or relationship. Some variables are not normally distributed and therefore do not meet the assumptions of parametric statistical tests. Using parametric statistical tests (such as a t-test or linear regression) on such data may give misleading results. In some cases, transforming the data will make it fit the assumptions better. For the stated reasons the data has been logarithm transformed so that the diversity of aircraft categories can be catered. By applying this transformation the data looks less skewed. The validation of transformation used can be performed through studying the correlation of different parameters untransformed and transformed in table 5. The improvement is correlation is evident after transformation for the propeller aircraft dataset. The same transformation has been done for the jet aircraft data sets. The improved correlation will help to build accurate surrogate models that will predict parameters with high confidence. The correlation matrix helps us to further study the dependence of response / dependent variable with predictor / independent variables. This matrix gives a deep insight in to how the models will be developed in further steps.

	Data condition	Wing span	MTOW	Power Av	Vmax	Range	Ceiling	Endurance
Wing span	Untransformed	1.000	0.889	0.417	0.169	0.662	0.689	0.761
	Transformed	1.000	0.897	0.820	0.452	0.644	0.717	0.811
MTOW	Untransformed	0.889	1.000	0.663	0.372	0.602	0.513	0.675
	Transformed	0.897	1.000	0.966	0.682	0.647	0.770	0.730
Power Av	Untransformed	0.417	0.663	1.000	0.590	0.289	0.398	0.276
	Transformed	0.820	0.966	1.000	0.729	0.581	0.786	0.645
Vmax	Untransformed	0.169	0.372	0.590	1.000	0.032	0.107	0.112
	Transformed	0.452	0.682	0.729	1.000	0.405	0.510	0.303
Range	Untransformed	0.662	0.602	0.289	0.032	1.000	0.571	0.556
	Transformed	0.644	0.647	0.581	0.405	1.000	0.666	0.729
Ceiling	Untransformed	0.689	0.513	0.398	0.107	0.571	1.000	0.574
	Transformed	0.717	0.770	0.786	0.510	0.666	1.000	0.723
Endurance	Untransformed	0.761	0.675	0.276	0.112	0.556	0.574	1.000
	Transformed	0.811	0.730	0.645	0.303	0.729	0.723	1.000

Table 5: Data transformation comparison via building correlation matrix

5.4 Model building

The least squares models were formed using MINITAB® V16. The models have been divided into two broad categories of propulsion i.e. propeller and turbojets/turbofans. It is pertinent to mention that multiple linear regression equations are verified using the diagnostic checks as mentioned in Chapter 3. The values of adjusted R^2 , total observations used in building model and F-test values with p-values are presented in Table 6 & 7.

Turbojet/Turbofan Aircrafts:

Parameters	Surrogate Model	Statistics Tests
Vmax	$LogV_{max} = 3.28 - 0.840 Logb + 0.420 LogT$	RSq(adj)=91%F=549.4 P<5%,Obs=128
Range	$LogRange = 2.13 + 0.197 Logb + 0.529 LogT$	RSq(adj)=77.6% F= 217.52 P<5%,Obs=130
Ceiling	$LogCeiling = 4.72 - 0.3 Logb + 0.169 LogS - 0.212 LogW + 0.216 LogT$	RSq(adj)=82.9%F=120 P<10%,Obs=125
Rate of Climb	$LogROC = 3.47 - 0.797 Logb - 0.668 LogS + 1.2 LogT$	R-Sq(adj)=95.4%F=307.95 P<5%,,Obs=55

Table 6: Surrogate models for Turbojet/Turbofan aircrafts

Propeller Aircrafts:

Parameters	Surrogate Model	Statistics Tests
Vmax	$LogV_{max} = 1.78 - 0.181 Logb + 0.258 LogP_a$	R-Sq(adj)=42.5%,F=18.37P<5%,Obs=67
Range	$LogRange = 0.848 + 0.664 LogW - 0.313 LogP_a$	R-Sq(adj)=83.2%,F=65.23P<10%,Obs=62
Ceiling	$LogCeiling = 3.12 + 0.469 Logb + 0.134 \log P_a$	R-Sq(adj)=54.7%,F=40.82P<5%,Obs=65
Rate of Climb	$LogROC = 2.49 - 0.621 Logb + 0.339 LogP_a$	R-Sq(adj) =47.6%,F=8.73P<5%,Obs=17
Endurance	$LogEndurance = -0.790 + 0.979 LogW - 0.422 LogP_a$	RSq(adj)=72.8%,F=74.80P<10%,Obs=72

Table 7: Surrogate models for Propeller aircrafts

5.5 Concluding remarks

Multiple variable models have been developed using multiple linear regression technique using method of least squares. The data collected has been logarithm transformed so as to reduce the skewness and make the data usable for linear regression. Further these models have passed diagnostics checks such as R^2, F test and t tests. Aircraft performance parameters such as range, maximum velocity, endurance, rate of climb and ceiling are predicted using geometric and propulsive properties.

CHAPTER 6

MODEL ACCURACY AND VALIDITY

6.1 Introduction

The model accuracy and validation is a critical step in order to build such surrogate models that can be used by design engineers. The validation process employed in this research is a three step process. These steps are explained as

i. The single variable models (using power law) and multiple variable models (using regression technique) are compared using quantitative criteria. The chosen model is then fed into second step.

ii. The chosen model is then compared with analytical equations of aircraft performance. In this step the analytical equations of maximum velocity, Endurance, service ceiling, Range and Rate of climb are compared with surrogate models so as to analyze the validity of models. The models predictor variables must also be a function of analytical equation, only then can model be considered valid.

iii. Lastly the model must also be able to predict those aircraft which are not part of the data set. For this step, a set of aircraft were chosen from all categories and there geometric parameters were fed to the models for validation.

6.1.1 Analysis of Models accuracy

Surrogate models have been developed using power laws and multiple linear regression models. Single variable models will be inadequate in order to predict aircraft performance due to inherent intricate aircraft designs. These single variable models are developed using power laws as data was found to be curvilinear. However after transforming data into linear framework multiple variable models were developed using regression technique. It is to be noted that by increasing variables into the model, it is not necessary that the models accuracy is increased. Individual models are developed and validated using above mentioned integrity test.

In this section we use quantitative criteria so as to further strength the model integrity. A possible one is the coefficient of determination R^2, which explains the percentage of total variation. R^2 also increases as the number of predictor variables used in the model increases even if these included variables are least significant. Hence, it is preferable to use the adjusted R^2 (equation mentioned in Chapter 3). In addition to coefficient of determination, the model performance can be evaluated using Mean Absolute Percentage error (MAPE), Root Mean Square Error (RMSE) and Nash-Sutcliffe (NSE). The Nash-Sutcliffe model efficiency coefficient (E) is commonly used to assess the predictive power of hydrological discharge models. However, it can also be used to quantitatively describe the accuracy of model outputs for other things than discharge. The equations for each criterion are as

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_{i_{calculated}} - y_{i_{observed}}}{y_{i_{observed}}} \times 100\right| \qquad (23)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{i_{calculated}} - y_{i_{observed}})^2} \qquad (24)$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(y_{i_{observed}} - y_{i_{calculated}})^2}{\sum_{i=1}^{n}(y_{i_{observed}} - \overline{y_{i_{observed}}})^2} \qquad (25)$$

Using these equations for models developed, a comprehensive comparion of models was performed and is displayed in Table 8. R^2,MAPE,RMSE and NSE are calculated for each model, R^2 is calculated in percentage and higher the percentage, model is fitted well. MAPE and RMSE should be closer to zero for a good fit. And NSE for a perfect fit for should be closer to 1. In this regard if the table 8 is studied in detail in respect of criterion values, multiple variable models are much fitted and explains the response variables much better. However single variable models are simpler in nature and are good for estimating design bounds.

Models	Category	Response	Predictors	R^2_{adj} %	MAPE	RMSE	NSE
Single variable models	Propeller	Endurance	b	79.13	136	7.344	0.531
	Turbojet/Propeller	V max	W	40.2	171.12	603	0.377
	Propeller	Range	Pa	95.31	251.3	826.4	0.715
	Turbojet	Range	T	82.28	40	2152.2	0.62
	Turbojet/Propeller	Range	W	86.56	254.4	1071	0.49
	Turbojet/Propeller	Ceiling	W	55.6	101.4	3835	0.45
	Turbojet/Propeller	Rate of climb	W	55.2	68.9	4041	0.55
Multiple variable models	Turbojet	Vmax	b,T	91	10.2	200	0.89
		Range	b,T	77.6	34.3	1238.5	0.85
		Ceiling	b,S,W,T	82.9	5.35	982	0.82
		Rate of Climb	b,S,T	95.4	17.30	1846.6	0.89
	Propeller	Vmax	b,Pa	42.5	11.65	14.8	0.38
		Range	W,Pa	83.2	14.8	15	0.76
		Ceiling	b,Pa	54.7	16.8	883	0.66
		Rate of Climb	b,Pa	47.6	24.3	110.6	0.33
		Endurance	W,Pa	72.8	29.15	5.34	0.70

Table 8: Models accuracy analysis

6.1.2 Validation of Selected model via Analytical Equations

The multiple variable models are then compared with the analytical equations of aircraft performance. The analytical functions are formulated using analytical equations of aircraft performance (already mentioned in chapter 3). It is highlighted that the statistical models must be comparable to the analytical equations, only then can a model be declared adequate in predicting the performance parameters of aircraft. To compare the analytical model with surrogate models, Table 9 can be helpful as a discrete comparison of model functions.

Parameters	Analytical equations function	Surrogate model function
Maximum Velocity	$V\max = f(T/Pa, W, S, Cdo, k, \rho\infty)$	$V\max = f(b, S, T/Pa)$
Range	$Range = f(V\infty, TSFC/SFC, \eta pr, L/D, Wo, W1)$	$Range = f(b, T, W)$
Ceiling	$Ceiling = f(W, S, T/Pa, \rho\infty, L/D, Cdo)$	$Ceiling = f(b, S, W, T/Pa)$
Rate of Climb	$ROC = f(W, S, T/Pa, \rho\infty, L/D, Cdo)$	$ROC = f(S, b, T/Pa)$
Endurance (for propeller only)	$Endurance = f(SFC, \eta pr, Wo, W1, Cl/Cd, \rho\infty, S)$	$Endurance = f(W, b)$

Table 9: Analytical models vs Surrogate models

From comparison of model functions, the surrogate models are in line with the analytical equation functions. For example if maximum velocity is a function of thrust/Power available, maximum take of weight, wing area, coefficient of drag at zero lift, k and density then the surrogate model developed is a function of wing span, Wing area and propulsive property. We now reach at a very interesting conclusion that we can build models with fewer parameters and still predicts the performance parameters to a high level of accuracy. Moreover time exhaustive calculations of Lift to drag ratios and coefficient of lift or drag in the initial phase of developing specification model can be avoided as prediction via geometric parameters still suffice for initial stages.

6.1.3 Prediction accuracy of Surrogate models

A set of aircraft data was collected in order to validate the surrogate models earlier formulated. The result of validation is mentioned in Table 10. As these models will help in the initial design process so the percentage deviation is satisfactory. With these initial values one can achieve the desired design with less number of iterations. Moreover the calculation via these models can also help build the foundations for design methods.

	% deviation in prediction				
Sample Aircraft	**Vmax**	**Range**	**Ceiling**	**Endurance**	**Rate of climb**
Dassault Ouragan	7	25.6	7.3		16
JH-7	10.2	5.3	0.005		-
F/A-18 Hornet	1.3	59.5	0.06		0.033
Airbus A300	18.3	26.4	0.08		-
Boeing 707-320	11.2	14.7	0.03		-
SIVA	20	24	18.7	6	17.6

Table 10: Validation result

6.2 Unique design emerged as outliers

During the scalability and regression analysis study, few unique designs of aircraft have emerged as potential outliers in the dataset. There can be multiple logical explanations for these designs being outliers, and vary from case to case.

i. The designs were high departure from conventional models. Such unique designs have then witnessed either high maintenance costs or frequent downtimes. Such overambitious designs were prematurely retired. Example for such designs are Concorde and Tu-144, they were supersonic passenger aircraft. Both these aircraft were a frequent outlier in model building.

ii. Then there are designs which are major technological advancements and have yet to see similar designs in production with subsequent service. A380 from airbus is only such design which is currently the biggest passenger carrier aircraft in-service. However other manufacturers are trying to build platforms similar to A380, after which it will not be a outlier anymore.

iii. Then there are some other unique or experimental design which have yet to see commercial use, however, there flight testing have been done. Centurion Helios and theseis, both UAVs are unique designs as they are wing alone design and have emerged as outliers.

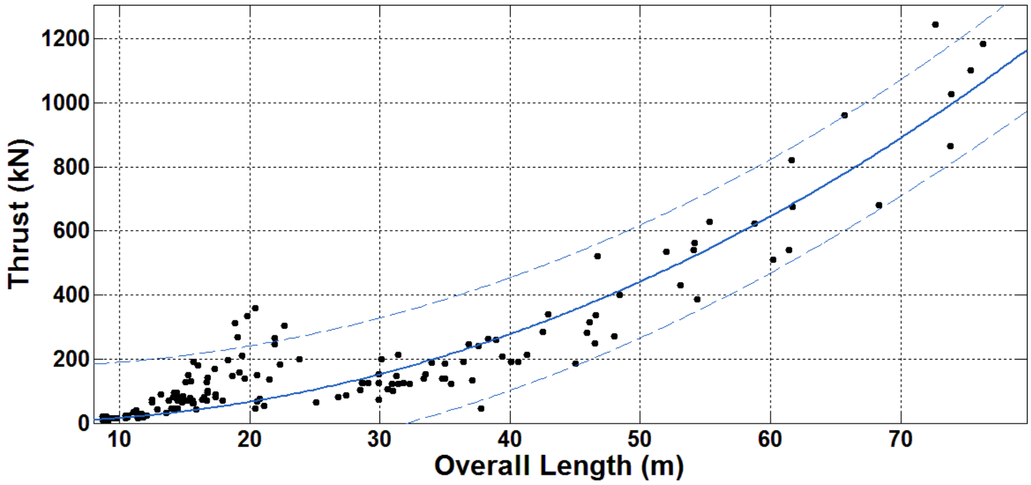
The study carried out also faced fewer limitations. As it’s a statistical study which depends on collected data set. If we increase the data set in terms of aircraft counts and different types, the statistical relations will be more refined. As aforementioned that the dataset was collected using Jane’s all the world aircraft 2007-2008[22], latest editions of Jane’s aircraft will include newer platforms and will also help the above mentioned outliers. In this regard it is to be noted that the surrogate models presented will need a constant update of data set in order to predict the newer and unique platforms of similar types with higher confidence

6.3 Unique design trends emerged during study

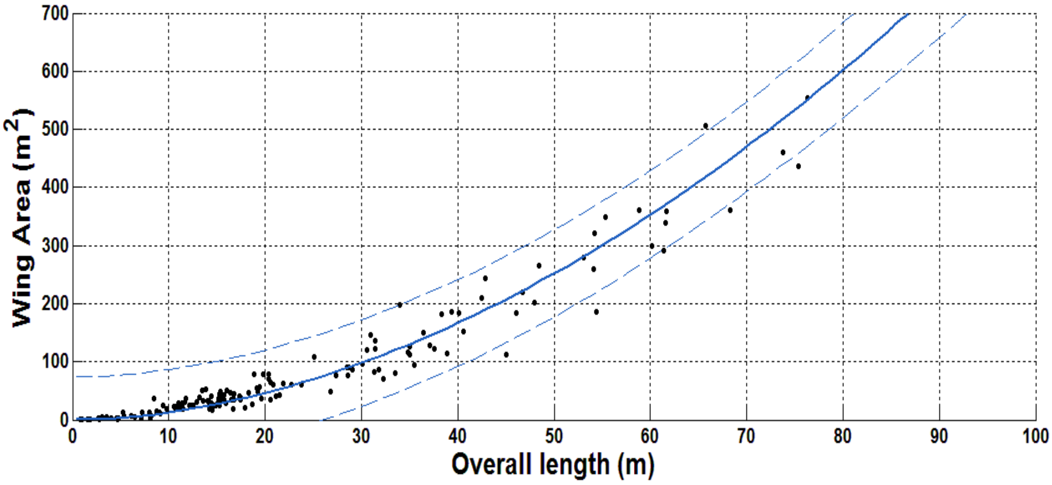
Further to that, while doing the scalability study certain unique relations were also observed, for which the analytical explanation was inadequate. However scalability exists and relations satisfy the data set with high degree of confidence. The overall length of the aircraft has explained several parameters such as weight, wing span, wing area, overall height, thrust, power available and range of aircraft. Further into the paper, analytical models for different performance parameters are shown in Table 11, where there is no dependence of overall length with performance parameters. But scalability exists with the overall length. This unique trend can be used in design estimations effectively. Few scalability graphs taken from table 11 are presented in figure 10.

Response (Y)	**Predictor(X)**	β	β**, 95% Cl**	**Adj R^2**	**normalization constant, α**	**Obs**
Weight	Length	2.069	(1.896, 2.242)	82.74	19.12	261
Wing Span	Length	1.026	(0.9813, 1.07)	92.49	0.7913	278
Wing Area	Length	1.709	(1.628, 1.79)	94.92	0.3025	163
Height	Length	0.9033	(0.864, 0.9425)	93.71	0.3881	196
Thrust	Length	2.082	(1.911, 2.252)	86.61	0.1284	148
Power	Length	5.533	(4.348, 6.719)	68.07	0.00079	71
Range	Length	1.241	(1.139, 1.343)	80.28	58.76	249

Table 11: unique design trends



a. Overall length vs Thrust with 95% confidence bounds



a. Overall length vs Wing area with 95% confidence bounds

Figure 10: Unique design trend graphs

CHAPTER 7

CONCLUSIONS & FUTURE WORK

7.1 Conclusion

This research forms the foundation of such techniques implementation in estimation of aircraft performance parameters.

i. The data for different aircraft parameters were collected using Jane journals and other online available data[22-24]. The collected data was scrutinized and consistency of units was maintained.

ii. Geometric and propulsive parameters were declared as predictors and aircraft performance parameters as response variables.

iii. Scalable relationships were formed using power laws as the data behavior was curvilinear. Single variable models are developed and results shows high confidence

iv. Multiple linear regression was used to build multiple variable models. The surrogate models presented in research predict aircraft performance to an adequate confidence level.

v. Single and multiple variable models are validated using a three point model accuracy and adequacy approach that is

   a. Analysis of model accuracy

   b. Validation of Selected model via Analytical Equations

   c. Prediction accuracy of surrogate models

vi. Results of three point validation shows that aircraft parameters are best fitted using multiple linear regression. But single variable model being the simplest provide low fidelity design bounds.

Surrogate models will in turn cut down the initial iterations to reach the specification model of design process. Once the specification model building time is cut down, it will complement the designers to focus on the computational models and verification/validation processes. Moreover these stated techniques will also be helpful for the strategic organization in estimating the adversary's aircraft to higher confidence level for further tailoring the counter strategies. Similar technique can be employed in other aviation design process whether it is building airfoil data or other systems such as propulsive and avionics system designs.

7.2 Future work

The future work that can be under taken in current research context is as under

i. The data set can be enhanced and surrogate modeling of estimated aircraft performance parameters can be refined.

ii. The technique employed in this research can be applied on other aviation areas. These areas can be broadly categorized as designer, manufacturing, maintenance and operational. Surrogate models can be developed in these areas. Such as estimating airfoil designs, calculating consumption and stock level of maintenance items and estimating different operational parameters of aircrafts.

iii. Further to multiple linear regression technique, nonlinear frame work study can be under taken so as to study large amount of aircraft data.

iv. As mentioned in chapter 4 section 4.5, unique design trends study may be further probed and investigated. These unique trends may tailor future design methodologies.

References

1. West, G.B., The Origin of Universal Scaling Laws in Biology. Elsevier, 1999. Physics A(263): p. P. 104-113.
2. Bettencourt, Growth, innovation, scaling, and the pace of life in cities. Proceedings of the National Academy of Sciences, 2007. 104(17): p. 7301-7306.
3. Moore, G.E., Cramming more components onto integrated circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp.114 ff. IEEE Solid-State Circuits Society Newsletter, 2006. 11(5): p. 33-35.
4. Stolzer, A.J., Fuel Consumption Modeling of a Transport Category Aircraft Using Flight Operations Quality Assurance Data: A Literature Review. Journal of Air Transportation, 2002. Vol. 7, No. 1 – 2002.
5. Henk Tennekes, The simple science of flight : from insects to jumbo jets. 2009.
6. Hebert, Use of Multifidelity and Surrogate Models in the Design and Development of Physics-Based Systems. Systems Engineering, 2016. 19(4): p. 375-391.
7. Hussain, Z., Estimation of flood quantiles at gauged and ungauged sites of the four major rivers of Punjab, Pakistan. Natural Hazards, 2017. 86(1): p. 107-123.
8. Almedeij, J., Modeling Pan Evaporation for Kuwait by Multiple Linear Regression. 2012. Volume 2012 p. 9 pages.
9. Josephat, P. And A. Ismail, A Logistic Regression Model of Customer Satisfaction of Airline. 2012, 2012. 2(4): p. 11.
10. Raj Bandyopadhyay, R.G. Predicting airline delays. 2012.
11. N.K. Taneja, A multi-regression analysis of airline indirect operating costs. 1968, massachusetts institute of technology flight transportation laboratory. P. 120.
12. Zhang, Surrogate-Based Optimization , Real-World Applications of Genetic Algorithms. Intech open science, 2012.
13. Han, Z.-H., S. Görtz, and R. Zimmermann, Improving variable-fidelity surrogate modeling via gradient-enhanced kriging and a generalized hybrid bridge function. Aerospace Science and Technology, 2013. 25(1): p. 177-189.
14. Kuya, Multifidelity Surrogate Modeling of Experimental and Computational Aerodynamic Data Sets. AIAA Journal, 2011. 49(2): p. 289-298.
15. Paul D. Gallaher, R.A.H., and Robert C. Williges, A regression approach to generate aircraft predictor information. 1978, Engineering physcology programs and aeronautics programs Office of Naval research.
16. Ghasemi Hamed, M., et al. Statistical prediction of aircraft trajectory : regression methods vs point-mass model. In ATM 2013, 10th USA/Europe Air Traffic Management Research and Development Seminar. 2013. Chicago, United States.
17. Liem, R.P., Multimission Fuel-Burn Minimization in Aircraft Design: A Surrogate-Modeling Approach, in Institute for Aerospace Studies 2015, University of Toronto: Canada. P. 160.
18. D. Allaire, K.W., Surrogate Modeling for Uncertainty Assessment with Application to Aviation Environmental System Models. AIAA JOURNAL 2010. Vol. 48 (No. 8).
19. Mack, Surrogate Model-Based Optimization Framework: A Case Study in Aerospace Design, in Evolutionary Computation in Dynamic and Uncertain Environments, S. Yang, Y.-S. Ong, and Y. Jin, Editors. 2007, Springer Berlin Heidelberg: Berlin, Heidelberg. P. 323-342.
20. Ashraf, A., Modeling of scalability laws for unmanned air vehicles, in Modeling of scalability laws for unmanned air vehicles. 2016, National University of Science and Technology: Islamabad. P. 66.
21. Anderson, J.D., Aircraft performance and design. International edition 1999.

22. Paul Jackson, K.M., Lindsay Peacock, Jane's All the aircraft world, K.M. Paul Jackson, Lindsay Peacock, Editor. 2007-2008.
23. Kenneth Munson, M.s., Jane's Unmanned aerial vehicles and targets, M.s. Kenneth Munson, Editor. 2005.
24. World UAV round up, in Supplement by American Institute of Aeronautics and Astronautics 2013.