AFRL-RQ-WP-TR-2022-0078

SURROGATE MODELING OF COMPUTATIONAL AERODYNAMIC RESPONSES FOR A GENERIC HYPERSONIC VEHICLE

Jacob R. Johanik[1], C. Corey Fischer[1], Benjamin J. Grier[2], Logan P. Riley[3], Ramana V. Grandhi[1], and José A. Camberos[3]

[1]**Air Force Institute of Technology**
[2]**The Ohio State University**
[3]**AFRL/RQHV, Vehicle Technology Branch**

MAY 2022
Interim Report

AIR FORCE RESEARCH LABORATORY
AEROSPACE SYSTEMS DIRECTORATE
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7542
AIR FORCE MATERIAL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

AFRL-RQ-WP-TR-2022-0078 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNED.//

JOSÉ A. CAMBEROS, PH.D., P.E.
Work Unit Manager
Vehicle Technology Branch

//SIGNED//

LT COL JOSEPH A. WAHLQUIST
Chief, Vehicle Technology Branch
High Speed Systems Division

//SIGNED//

TIMOTHY A. EYMANN, PH.D.
Technical Advisor, Vehicle Technology Branch
High Speed Systems Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

1. REPORT DATE *(DD–MM–YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From — To)*
MAY 2022	Final Report	September 2020 - March 2022

4. TITLE AND SUBTITLE
Surrogate Modeling of Computational Aerodynamic Responses for a Generic Hypersonic Vehicle

5a. CONTRACT NUMBER
N/A

5b. GRANT NUMBER
N/A

5c. PROGRAM ELEMENT NUMBER
N/A

6. AUTHOR(S)
Jacob R. Johanik, C. Corey Fischer, Benjamin J. Grier, Logan P. Riley, Ramana V. Grandhi, and José A. Camberos

5d. PROJECT NUMBER
N/A

5e. TASK NUMBER
N/A

5f. WORK UNIT NUMBER
N/A Q1TM

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology
Wright-Patterson Air Force Base, OH 45433-7542
United States Air Force

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT-ENY-MS-22-M-295

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
Aerospace Systems Directorate
Wright-Patterson Air Force Base, OH 45433-7542
Air Force Material Command
United States Air Force

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/RQHV

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-RQ-WP-TR-2022-0078

12. DISTRIBUTION / AVAILABILITY STATEMENT
DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited; 88ABW-2022-0190; Cleared 07/18/2022.

13. SUPPLEMENTARY NOTES
This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT
In the field of multidisciplinary hypersonic vehicle design, striking the balance between the accuracy and efficiency of a predictive aerodynamic response model is a significant challenge. In response to this challenge, the objective of this research is to evaluate the aerodynamic performance of a Generic Hypersonic Vehicle GHV using the technique of surrogate modeling Computational Fluid Dynamics CFD data points across a large range of flight conditions. To accomplish this, the full CFD process was conducted by preparing the vehicle geometry, generating a grid, computing the flow, and post-processing the data. A three-dimensional, quasi-random distribution of 1000 points defined the design space of the study which consisted of varied Mach number, angle of attack, and flight altitude. Using inviscid CFD training data from the design space, surrogate models of integrated forces and critical surface pressures were generated using the Kriging method, and the suitability of these models was evaluated using additional validation CFD data. Additional studies were conducted to evaluate the optimal correlation and regression functions for the Kriging models and to determine the optimal number of training points needed for a sufficiently accurate model.

15. SUBJECT TERMS
Surrogate, Meta, Model, Reduced Rrder Model (ROM), Computational Fluid Dynamics (CFD), Euler, Kriging, Gaussian Process, Regression

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			José A. Camberos
Unclassified	Unclassified	Unclassified	Unclassified	103	**19b. TELEPHONE NUMBER** *(include area code)* (937) 713-7055

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

LIST OF ACRONYMS

AGARD : Advisory Group for Aerospace Research and Development
BLUP : Best Linear Unbiased Prediction
CAD : Computer-Aided Design
CBAERO : Configuration-Based AERODynamics
CFD : Computational Fluid Dynamics
CFL : Courant–Friedrichs–Lewy
CPU : Central Processing Unit
DACE : Design and Analysis of Computer Experiments
DES : Detached Eddy Simulation
DNS : Direct Numerical Simulation
ERT : Eckert Reference Temperature
ESP : Engineering Sketch Pad
FEA : Finite Element Analysis
GHV : Generic Hypersonic Vehicle
HPC : High Performance Computing
ITAR : International Traffic in Arms Regulations
LES : Large Eddy Simulation
MLE : Maximum Likelihood Estimation
NASP : National Aero-Space Plane
NRMSE : Normalized Root Mean Squared Error
PDE : Partial Differential Equation
RANS : Reynolds Averaged Navier Stokes
RMSE : Root Mean Squared Error
ROM : Reduced Order Model
TC : Test Condition
WAC : Without Attitude Control

1 SUMMARY

In the field of multidisciplinary hypersonic vehicle design, striking the balance between the accuracy and efficiency of a predictive aerodynamic response model is a significant challenge. In response to this challenge, the objective of this research is to evaluate the aerodynamic performance of a Generic Hypersonic Vehicle GHV using the technique of surrogate modeling Computational Fluid Dynamics CFD data points across a large range of flight conditions. To accomplish this, the full CFD process was conducted by preparing the vehicle geometry, generating a grid, computing the flow, and post-processing the data. A three-dimensional, quasi-random distribution of 1000 points defined the design space of the study which consisted of varied Mach number, angle of attack, and flight altitude. Using inviscid CFD training data from the design space, surrogate models of integrated forces and critical surface pressures were generated using the Kriging method, and the suitability of these models was evaluated using additional validation CFD data. Additional studies were conducted to evaluate the optimal correlation and regression functions for the Kriging models and to determine the optimal number of training points needed for a sufficiently accurate model.

2 INTRODUCTION

2.1 Motivation

In the field of hypersonic vehicle design, there are numerous difficulties within the engineering design process because of the non-linearity [1] and strong coupling of aerothermalelastic effects of hypersonic flow [2]. Because hypersonic flow is highly energetic while flying around and above Mach 4, unique phenomena occur such as large surface heat fluxes, shock/boundary-layer interactions, low-density flow, and shock-shock interaction among others [3]. These phenomena need to be well understood and predicted when designing a hypersonic vehicle in order to avoid unexpected vehicle or mission failure. However, the issue with predicting these hypersonic phenomena is that it requires extensive prior knowledge of the vehicle geometry and a considerable amount of time and money using accurate, high-fidelity models.

Since the engineering process is iterative, the amount of time and money used for every preliminary design iteration should be minimized. For the preliminary design stage of the engineering process, predictions using high-fidelity models, *i.e.* RANS, DES, or DNS, are rarely used because of the large costs required to complete them. High-fidelity models may provide accurate predictions of the vehicle, but they are not needed to provide a general understanding of the vehicle performance used in preliminary design. Low-fidelity models, on the other hand, are not as accurate as high-fidelity models, but they require less computational time and money. Using low-fidelity models for the preliminary design of a hypersonic vehicle is desirable because of the low cost. Although the accuracy of low-fidelity models is lower, they still provide general aerothermal predictions of the vehicle response that can be used to further improve the design.

In the discipline of hypersonic aerothermodynamics, computational predictions are typically preferred over ground testing and flight testing because of its relative low cost and ease of implementation. Hypersonic ground tests are very costly to run because of the high Mach and large temperatures needed in a confined environment; and hypersonic flight testing incurs massive costs due to regulations and material. Computational Fluid Dynamics (CFD) only requires a well-equipped computer or cluster, and thus it is typically the preferred tool when predicting aerothermodynamics in the preliminary design phase. The field of CFD contains both low-fidelity models, such as panel methods, along with higher-fidelity models such as Reynolds Averaged Navier-Stokes (RANS) simulations.

Thus, when exclusively considering the preliminary design phase of a hypersonic vehicle, a low-fidelity method of providing quick results is needed. One particular method to accomplish this is the Reduced Order Model (ROM), also referred to as a surrogate model. The philosophy behind a surrogate model is to utilize a limited number of expensive samples from an unknown function within a design space to develop an interpolation that estimates the function values anywhere inside that design space [4]. For example, a surrogate may be able to quickly predict the drag that a vehicle produces across a range of input parameters (Mach, altitude, etc.). This model is valuable when the underlying function describing the drag of the vehicle is unknowable. The main drawback is that a sizable number of data points is needed to train a surrogate model which can be expensive to obtain. Surrogates are a powerful tool in the field of hypersonic vehicle aerothermodynamics because it has the capability of predicting non-linear behaviors that are

common in hypersonic flow.

2.2 Research Objectives

The primary objective of this research was to develop surrogate models of key aerothermodynamic parameters for a Generic Hypersonic Vehicle (GHV) (Figure 1) across a design space of varying flight conditions. The purpose of these surrogate models was to rapidly predict certain aerodynamic responses of the vehicle across a range of design parameters. The design parameters to be analyzed were the Mach number, angle of attack, and the flight altitude. The aerodynamic responses that were computed were the lift, drag, and pitch moment of the entire aircraft, then surface pressures and surface heat fluxes at certain critical points on the vehicle surface.

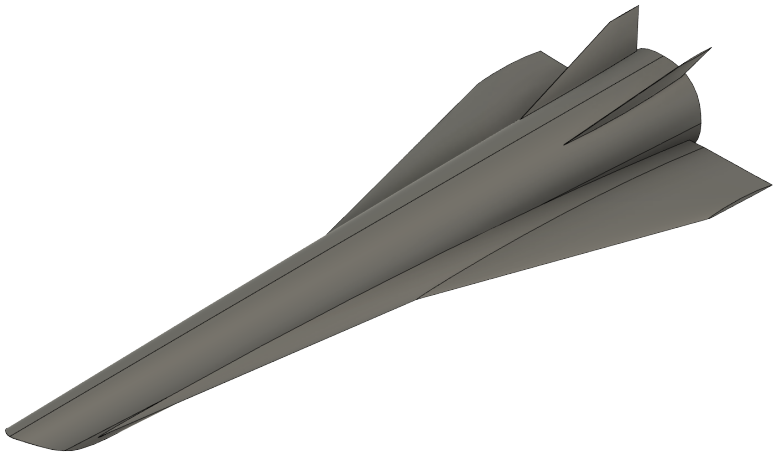


Figure 1: Isometric View of the GHV

The surrogate models required data points from CFD in order to predict the aerodynamic responses across the design space. This data was collected using an Euler (inviscid) CFD code. This work was intended to be used within the preliminary and conceptual design phases of vehicle design, thus a low to mid fidelity CFD code was desired for aerodynamic predictions. Since the vehicle geometry and mission is subject to change within these early design phases, it was assumed that Euler was the highest fidelity to be used for surrogate model predictions in the preliminary/conceptual design stage. Using higher fidelity models (*i.e.* RANS) would incur significantly more cost in time, money, and computational resources that are unreasonable for this design phase.

The surrogate models of each aerodynamic response will be inspected to determine the optimal kriging correlation and regression function. Additionally, the number of training points for an

accurate surrogate of each response will be determined. Trends of the responses with respect to the design space will be inspected.

The three design parameters, Mach, angle of attack, and altitude, were chosen for the purpose of describing the aerodynamic behavior for different flight conditions. Other potential parameters could have included the sideslip angle, control surface deflections, etc. However, for the scope of this research, the three aforementioned parameters were used to reduce complexity. The bounds on the design parameters are specified in Table 1. The baseline condition represented the cruise condition of the GHV and the min and max bounds were chosen to contain possible conditions the GHV could experience during a typical flight.

Table 1: Sample Design Space

Parameter	Min	Baseline	Max
M	3.0	6.0	8.0
α (°)	-4.0	0.0	6.0
h (m)	15,000	25,000	27,500

2.3 Document Overview

In this research report, a thorough literature review was conducted in Chapter 3. Journal articles and reports about the GHV geometry, CFD, and Kriging analysis are discussed. The research methodology used in this report is discussed in Chapter 4 where the CFD process and surrogate modeling is reviewed. After discussing the process of CFD data collection, the data is analyzed and modeled in Chapter 5. Finally, the research findings and future work are presented in Chapter 6.

3 BACKGROUND AND LITERATURE REVIEW

The second chapter includes a review of published literature and background theory that were pertinent to the research conducted in this study. A brief overview of notable hypersonic vehicles and issues pertaining to those hypersonic vehicles will be discussed. Surrogate modeling history and techniques will be introduced to provide adequate context to the analysis methods used during this research. Additionally, the origins of the GHV and its design features will be discussed.

3.1 Hypersonic Introduction

3.1.1 A Brief Hypersonics History

3.1.1.1 Re-entry Capsules

The realm of hypersonic flight was initially reached by the power of rocket motors. The first man-made object to reach hypersonic speeds was the WAC corporal rocket mounted on a V-2 lower stage in February 1949. This two-stage configuration was launched from the White Sands Proving Ground in New Mexico and the small, slim WAC corporal rocket reached a maximum speed of 5150 mph and a maximum altitude of 244 miles [3]. The multi-stage rocket configuration introduced by the WAC corporal was improved and used in subsequent years to propel spacecraft out of the atmosphere into space.

The first man to enter the hypersonic regime was U.S.S.R. Flight Major Yuri Gagarin as he reentered the atmosphere from the first manned orbit around the Earth in April 1961. Gagarin piloted the Vostok 1 spacecraft which orbited once around Earth and reentered the atmosphere with a speed above Mach 25 [3]. The capsule received damage and was charred from the extreme conditions. Later that year, the U.S. accomplished the second manned flight to space when astronaut Alan B. Shepard flew on the Mercury-Redstone 3 rocket on a suborbital trajectory with a peak speed near Mach 5. The Mercury-Redstone launch vehicle was a single stage rocket initially designed as a tactical ballistic missile [5]. The rocket was modified to be “man-rated” in order to safely launch a pilot to space. The Vostok and Mercury programs proved that single and multi-staged rockets were able to propel blunt capsules up to hypersonic speeds and even to Earth orbit, but the capsules had a distinct lack of rigorous control.

3.1.1.2 Rocket Powered Aircraft

A more controllable alternative to the multi-stage rocket and capsule configuration is the rocket powered aircraft. An early example of this configuration is the Bell X-1 which featured a traditional aircraft shape powered by a rocket motor using liquid oxygen and alcohol fuel [6]. On October 14, 1947, Capt. Chuck Yeager flew the X-1 past the sound barrier for the first time in history as he reached a speed of Mach 1.06 and an altitude of 43,000 ft. The X-1 was dropped from a B-29 at 25,000 ft so that it could reach supersonic speeds at the target altitude. The rocket system used by the X-1 was only able to provide thrust for a maximum of about 2.5 minutes [7]. During the X-1 flight tests, a drop in aircraft controllability was noticed while flying transonic due to the flow compressibility and other issues such as buffeting were observed. The first of the X planes opened the door to further supersonic aircraft research.

In the late 1950's, scientists discovered how hypersonic flow was drastically different from supersonic flow, and the need to characterize it grew as the US space program grew in priority. This need was answered using the X-15 hypersonic research vehicle. The X-15 mission evaluated and predicted the extreme conditions experienced during hypersonic flight. The X-15 was capable of flying up to Mach 6.7 and altitudes up to 250,000 ft using a throttleable liquid fuel rocket that provided a maximum 57,000 pounds of thrus . The X-15 only had fuel to provide thrust for approximately 90 seconds after dropping from the B-52 mothership. This was the limiting factor to its top speed and altitude during flight despite the fact that the vehicle's volume was dominated by its fuel tanks. The X-15 provided extremely useful data describing the hypersonic regime which has been used and cited up to the present [8].

3.1.1.3 Scramjet Power

Rocket power allowed the X-1 and X-15 to perform their groundbreaking, high-speed flights, but these vehicles could only sustain thrust for a few minutes after launch. The primary limitation of rocket power is that the vehicle is required to carry oxidizer internally in order to burn the fuel [9]. Rocket motors have a lower specific impulse (Isp) than turbojet engines because of the additional weight of the oxidizer. The specific impulse is a metric for motor efficiency and is defined as the amount of thrust per unit weight of propellent burned. A typical turbojet has an Isp around 6000 seconds while a typical rocket may have an Isp of 350 seconds at sea level. However, at high altitudes and hypersonic speeds, turbojet engines are not capable of maintaining sufficient thrust for sustained flight. Rockets are capable of accelerating vehicles to hypersonic speeds, but they do so very inefficiently. In order to obtain efficient hypersonic flight, a ramjet or scramjet is needed to propel the vehicle because of the limitations to rockets and turbojets.

A supersonic ramjet, or scramjet, operates by compressing air via a ramp inlet into a compression chamber where fuel is injected and combusted. The internal pressure is provided solely from the incoming air due to the very high stagnation pressures of the hypersonic flow. Scramjets maintain supersonic flow during the entire cycle to reduce the overall drag. Ramjets on the other hand have larger drag above Mach 6 because they slow the hypersonic flow down to subsonic speeds for combustion. Both scramjets and ramjets are desirable motors for highly supersonic to hypersonic vehicles because they do not contain any moving parts that compress or extract energy from the internal flow. Scramjets are comprised of four main components which are shown in Figure 2:

1. *Inlet* - The inlet is responsible for compressing the oncoming hypersonic flow by using a series of oblique shock waves. There are many different types of inlet configurations, but they all utilize shock waves for compression.
2. *Isolator* - The isolator allows the supersonic airflow to compress more through the use of a "shock train", a series of reflected internal oblique shocks. The isolator also provides stability for the scramjet cycle during operation by preventing flameouts.
3. *Combustor* - The combustor stage injects fuel into the airflow and ignites it via flameholders. The combustor adds energy into the airflow which is used to provide thrust to the vehicle.

4. *Nozzle* - The nozzle expands the high-energy, supersonic gas from the combustor and accelerates it out the rear of the vehicle. The net acceleration of the gas exiting the nozzle generates thrust.

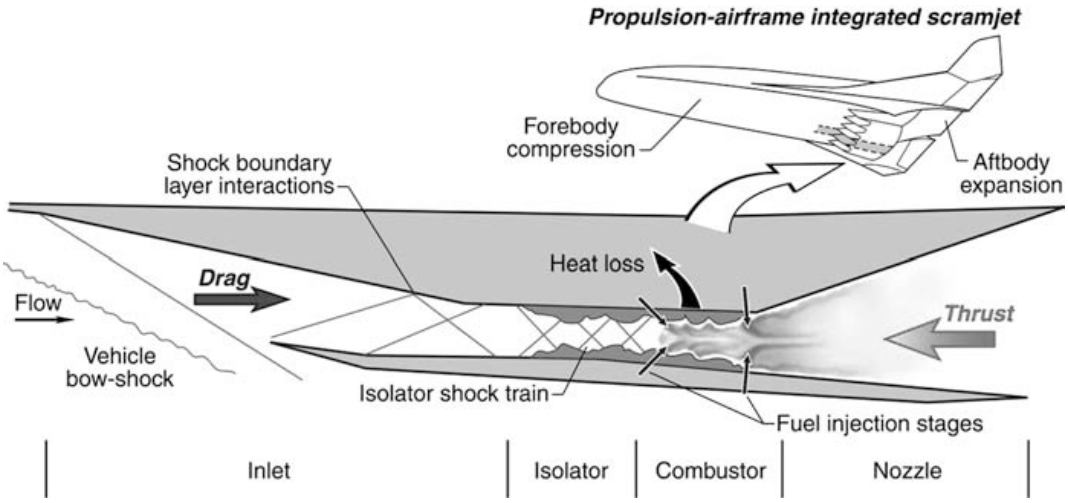


Figure 2: Illustration of a Typical Scramjet Configuration [9]

One of the first earnest attempts to implement a scramjet in a hypersonic vehicle is the National Aero-Space Plane (NASP) program (or X-30) which started in 1986 and ended in 1993 and was a joint program of the Department of Defense (DoD) and National Aeronautics and Space Administration (NASA). The ambitious NASP program sought to develop and evaluate the feasibility of a manned, hypersonic vehicle that could launch from the ground and accelerate into low-Earth orbit in a single-stage [10]. The NASP intended to utilize a combined cycle of a hydrogen fuel scramjet and a rocket motor so the vehicle could launch vertically from the ground, accelerate to hypersonic speeds, and reach Earth orbit all without using any booster stages. Figure 3 shows a wind tunnel model of the NASP. Valuable research of hypersonic propulsion was conducted during the NASP program, but the program was too ambitious because of the lack of technology required to meet all the design requirements. During the last year of development, the NASP design weight grew out of proportion due to lack of engine performance and high drag. Because of this, the program eventually lost funding and terminated in 1993.

Shortly after the NASP program ended, the need for a functional hypersonic vehicle demonstrator arose and the Hypersonic Experiment (Hyper-X) or X-43 program began. NASA specified three objectives for this new program: first, data was needed on a flight-tested scramjet; secondly, the technology and design tools developed from the NASP program needed to be validated; and lastly, the program needed to meet the strict goals of the new ”faster, better, cheaper” concept set by NASA Administrator, Daniel S. Goldin [9]. The X-43 was designed to operate at Mach 7 or Mach 10 at an altitude of around 94,000 ft. The research vehicle was propelled to cruise conditions using a modified Pegasus solid fuel rocket booster dropped from a B-52 mothership. The X-43

Figure 3: NASP Wind Tunnel Model [9]

scramjet was powered using gaseous hydrogen which was notoriously difficult to handle due to its very low density and volatility. The Hyper-X program relied on new predictive tools such as Monte Carlo simulations and CFD to predict the extreme hypersonic flight conditions.

There were three X-43 test flights in total. The first flight on June 2001 was a failure due to the Pegasus booster flight control system malfunctioning shortly after release from the mothership. A rigorous mishap investigation was conducted and modifications were made to the booster. The second flight occurred on March 2004 which flew successfully up to a max speed of Mach 6.83. The hydrogen scramjet operated successfully and was able to accelerate the vehicle for a short 10 seconds. The vehicle gathered more data by performing several hypersonic maneuvers after the engine burned out. The vehicle and booster splashed into the Pacific after the flight and were not recovered. The third X-43 flight on November 2004 pushed the envelope further and achieved a top speed of Mach 9.6. The scramjet operated for 11 seconds total and the vehicle performed additional maneuvers similar to the second flight. The Hyper-X program proved that scramjet

technology was feasible and this opened the door to future development of more advanced scramjets.

3.1.2 Hypersonics Issues

The realm of hypersonics is characterized by a large number of significant engineering challenges that need to be addressed when designing a hypersonic vehicle. The historical examples above each experienced unique issues related to hypersonics that required novel vehicle design choices. As the Mach number grows larger, certain hypersonic phenomena begin to dominate. John Anderson in his book *Hypersonic and High-Temperature Gas Dynamics* [3] defines these hypersonic phenomena and are summarized as follows:

- *Thin Shock Layers* - Oblique shock theory states that as the Mach number increases, the shock wave angle grows smaller relative to centerline. At very high Mach numbers, the distance between the shock waves and the vehicle body are very low. This is commonly referred to as a ”Shock Layer”. The thin shock layers experienced during hypersonic flight lead to many other effects which will be discussed.
- *Entropy Layer* - The majority of hypersonic vehicles feature a blunt nose or blunt body which induces a bow shock in front of the vehicle due to the extremely high pressures near the stagnation point. The bow shock is nearly normal (strong) around the stagnation point and more oblique (weak) around the side of the body as illustrated in Figure 4. According to thermodynamic theory, the entropy generated from a shock wave is greater across a strong shock opposed to a weak shock. Because of the curvature of the bow shock and the high energy flow, a prominent entropy gradient is formed aft of the shock which propagates downstream. This effect is called the ”entropy layer”. This entropy layer interacts noticeably with the viscous boundary layer, and the entropy layer induces a large amount of vorticity due to Crocco’s theorem.
- Viscous Interaction - A hypersonic boundary layer creates a cascade of effects resulting from air viscosity and the high-speed flow. In a hypersonic boundary layer, the high-speed flow is significantly slowed down which converts most of the kinetic energy of the flow to internal energy. The increase of internal energy results in higher gas temperature which causes an increase of air viscosity and lower density. The low density and high viscosity causes the boundary layer to grow quickly, increasing the thickness. Since thin shock layers are prevalent in hypersonic flow, the viscous boundary layer commonly interacts with the inviscid shock waves creating additional effects such as increased internal energy and in turn, increased boundary layer thickness. All in all, the viscous-inviscid interactions cause hotter air and thicker boundary layers, and this additional heat must be accounted in the design of a hypersonic vehicle.
- *High-Temperature Flows* - As discussed before, the high-energy flow and large friction in a hypersonic boundary layer generates very high temperatures. As the airflow grows hotter, the fundamental properties of the flow begin to change. The flow may start to dissociate,

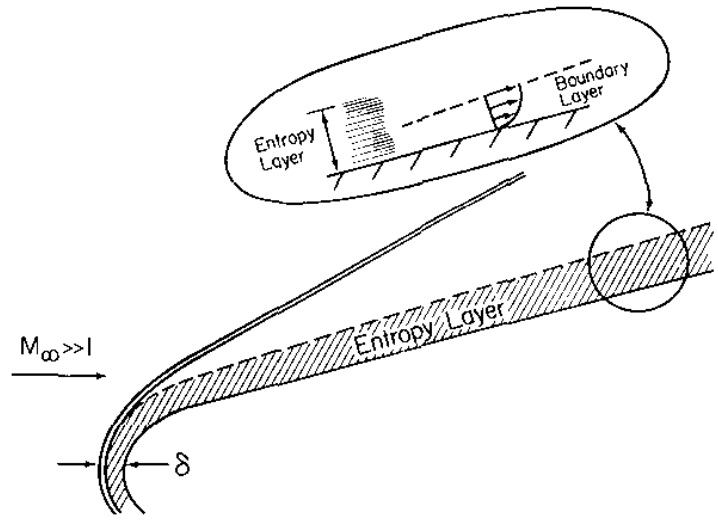


Figure 4: Entropy Layer Illustration [11]

chemically react, vibrate, radiate, and ionize depending on the temperature. These phenomena, called "high-temperature effects", are not modeled within traditional thermodynamics, thus the perfect gas law, calorically perfect gas assumption, and equilibrium assumption no longer apply and new models need to be considered. High-temperature effects may cause a series of unique issues such as force/moment disturbances, communication blackouts due to ionization, and non-equilibrium flow which is notoriously difficult to model.

- *Low-Density Flow* - Hypersonic vehicles fly at very high speeds, upwards of 4,000 feet per second. At these speeds, the dynamic pressure, defined in Equation (1), is dominated by the V^2 term resulting in very high pressure.

$$q = \frac{1}{2}\rho V^2 \tag{1}$$

Because of this, most hypersonic vehicles operate at high altitudes where the density, ρ, is smaller to reduce the dynamic pressure to suitable flight conditions for the vehicle's application. As the air density is lowered, air molecules spread further apart, increasing the mean free path and lowering the molecule collision rate. In low-density flow, the air can no longer be treated as a continuum of particles, but rather as a rarefied flow where individual air particles must be considered. Kinetic theory is used to describe the behavior of rarefied gas and hypersonics and is primarily reliant on the quantum mechanical properties described by the Boltzmann equation.

Each of these issues introduce increased complexities with the prediction of hypersonic vehicle performance. The non-linear behavior of the aero loads and surface heating can raise difficulties

within the preliminary and conceptual design of a hypersonic vehicle. Small changes in flight conditions could potentially affect the heating loads greatly, and these effects need to be modeled and understood to properly design a working hypersonic vehicle. CFD has the capability of predicting these non-linear effects at a desired fidelity, but at a large computational cost. The method of kriging using CFD data will be utilized to adequately model these hypersonic non-linear issues.

3.2 CFD Code Types

Within the discipline of CFD, there are multiple levels of fidelity that different codes can perform. Lower fidelity code typically makes many assumptions on the behavior of the flow which can overlook certain features. High fidelity code makes fewer assumptions and is far more accurate, but it comes with the additional expense of computational cost. Starting from low fidelity to high fidelity, common CFD codes are listed and briefly described below.

- *Panel Methods using Potential Flow* - This approach models 3D geometries using a discrete number of surface panels, also called singularities, each with unknown strengths. The most basic implementation of a panel method uses the Laplace potential flow theory to define the flowfield surrounding the vehicle using these singularities. The general Laplace PDE used for potential flow theory is,

$$\mathbf{\Phi}_{xx} + \mathbf{\Phi}_{yy} + \mathbf{\Phi}_{zz} = \nabla^2 \mathbf{\Phi} = 0 \quad [12] \tag{2}$$

  Panel methods assume the flow is inviscid which neglects any effects such as boundary layers and turbulence. Additionally, potential flow theory is only valid for subsonic flow, making it useless for a hypersonic application. Corrections and modifications to the basic implementation have been demonstrated to model supersonic flow [12].

- *Panel Methods using Newtonian Theory* - Again, this approach models a 3D geometry using surface panels. However, only the surface flow conditions are computed and not the surrounding flowfield because of the implementation of Newtonian flow theory [12]. This theory only considers the incidence angle of a given panel with respect to the oncoming flow to calculate the pressure on the panel. The primary equation for Newtonian theory is,

$$C_p = 2\sin^2\theta \quad [3] \tag{3}$$

  where C_p is the local pressure coefficient and θ is the local panel/wall incidence angle. Newtonian theory is an inviscid method that is only applicable to highly supersonic and hypersonic flow regimes. CBAERO is an example of this type of panel method [13].

- *Euler* - This CFD code utilizes the finite volume method to compute the flow properties of a domain. This method requires a 3D discretized computational domain, *i.e.* a grid, to

operate. Euler CFD uses the Navier-Stokes equations with the assumption that the flow is inviscid. Compressibility of the flow is simulated with Euler. The three governing PDEs for each dimension are,

$$\begin{aligned} \rho\frac{Du}{Dt} &= -\frac{\partial\rho}{\partial x} \\ \rho\frac{Dv}{Dt} &= -\frac{\partial\rho}{\partial y} \\ \rho\frac{Dw}{Dt} &= -\frac{\partial\rho}{\partial z} \end{aligned} \tag{4}$$

where u, v, and w are the velocity components; ρ is the density; x, y, and z are the positional components; and t is time [12]. The Euler code has the capability of modeling both subsonic and supersonic flow.

- *RANS* - Reynolds Averaged Navier-Stokes (RANS) utilizes the full Navier-Stokes equation set which simulates viscous shear of the flow. The solution calculated using RANS is time-averaged, meaning that any unsteady effects from turbulence is neglected. The 2D Reynolds averaged momentum equation for the x-direction is given as,

$$\rho u\frac{\partial u}{\partial x} + \rho v\frac{\partial v}{\partial y} = -\frac{\partial p}{\partial x} + \frac{\partial}{\partial y}\left[\mu\frac{\partial u}{\partial y} - \rho\bar{u'v'}\right] \quad [12] \tag{5}$$

where the $\rho\bar{u'v'}$ term is the Reynolds stress which accounts for transfers of momentum due to turbulent fluctuations [12]. This term is estimated using a turbulence model which relies on empirical data. RANS is unable to model any eddies or unsteady turbulence features of any scale.

- *Large Eddy Simulation (LES)* - The LES method increases the fidelity from RANS by computing larger, higher energy turbulent structures while estimating smaller scale turbulence structures. This is accomplished by a spacial filter that neglects any small scale flow effects while allowing the computation of the more significant, larger features. LES is a compromise between RANS and Direct Numerical Simulation (DNS) which resolves all turbulence scales. LES is still an extremely computationally expensive method. [12]
- *Detached Eddy Simulation (DES)* - The DES method is a hybrid code that both uses RANS and LES for different regions of the flow. The typically steady boundary layer is modeled using RANS due to its computational efficiency and lower grid requirements, and LES is used outside the boundary layer to resolve detached eddies from flow separation. A switching function based on the spacial distance to the nearest wall is used to switch between RANS near the wall and LES further away from the wall. [12]
- *DNS* - DNS fully resolves all turbulent scales by solving the full Navier-Stokes equations. This requires extremely highly resolved grids so the smallest turbulence scales can be

resolved. DNS is only feasible for small, simple geometries due to the limitations of today's computers. [12]

This research utilized an Euler code type as it was assumed to be the highest justifiable fidelity for use within the preliminary/conceptual design phase. Using a higher fidelity model may result in a more accurate flow response; however, given the intention of this work in constructing surrogate models for use in a conceptual design phase, the added accuracy may be negated by the inherent error associated with utization of surrogate models and thus the additional costs in time, money, and computational resources are not justified.

3.3 Eckert Reference Temperature Method

The convective heat transfer on the surface of a hypersonic vehicle is a pertinent parameter to consider when designing a hypersonic vehicle. The structural and thermal management systems of a hypersonic vehicle require aero-heating load predictions in order to properly decide on which materials to use and where they should be placed. However, predicting the surface heat flux is a complex problem [14].

To calculate the surface heat flux and convective heat transfer values, the Eckert Reference Temperature (ERT) method is used [15, 16]. This semi-empirical method uses boundary-layer relations from incompressible flow theory and corrects for compressibility using thermodynamic properties of air [17]. Essentially, ERT uses known, local pressure, temperature, Mach, and air velocity values on the edge of the boundary layer and estimates the convective heat transfer coefficient of the wall. The surface heat flux can then be calculated if the wall temperature is known. The formulation of ERT follows the process discussed in the dissertation, *Reduced-Order Aerothermoelastic Analysis of Hypersonic Vehicle Structures* by Falkiewicz [17].

The convective heat flux, $\dot{q}$ on the surface of the vehicle is defined as,

$$\dot{q} = h(T_w - T_r) \tag{6}$$

where h is the convective heat transfer coefficient, T_w is the wall temperature, and T_r is the recovery temperature. Since the viscous force responsible for slowing down the air is irreversible, and not all the kinetic energy is converted to thermal energy, a recovery factor (r_f) is defined:

$$r_f = \frac{T_r - T_e}{T_t - T_e} \tag{7}$$

where T_e is the air temperature at the boundary layer edge and T_t is the total temperature.

Note that the properties at the edge of the boundary layer are evaluated using results from Euler CFD. Since the Euler equation set is an inviscid model, boundary layer effects are not

implemented. Thus, the surface parameters (T_{surf}, P_{surf}, M_{surf}, and V_{surf}) calculated using Euler CFD are assumed to be equal to the boundary layer parameters (T_e, P_e, M_e, and V_e) used in ERT.

Using the local boundary-layer edge temperature and Mach, the total temperature can be calculated using the isentropic relationship,

$$T_t = T_e \left[1 + \frac{\gamma - 1}{2} M_e^2\right] \tag{8}$$

where γ = 1.4. The recovery factor for turbulent and laminar flow is calculated using the Prantl number,

$$r_f = Pr^{1/2} \quad \text{(laminar)} \tag{9}$$

$$r_f = Pr^{1/3} \quad \text{(turbulent)} \tag{10}$$

The Prantl number is assumed to be constant at the value $Pr = 0.7$. Since the flight conditions being analyzed will be have a Mach number greater than 3, the flow will be assumed to be turbulent across the surface because the Reynolds number for the range of flight conditions were each well over 500,000 across the entire vehicle. Thus, Equation (10) will be used. Because of the high-speed flow being modeled, the airflow properties near the boundary layer are subject to significant change. Eckert introduced a reference temperature, T^*, at which the boundary layer air properties are evaluated.

$$T^* = T_e + 0.50(T_w - Te) + 0.22(T_r - T_e) \tag{11}$$

Using the reference temperature, the reference viscosity, μ^*, and reference density, ρ^*, are calculated using Sutherland's Law and the ideal gas law, respectively:

$$\mu^* = \frac{C_1 T^{*3/2}}{T^* + C_2} \tag{12}$$

$$\rho^* = \frac{P_e}{RT^*} \tag{13}$$

Where $C_1 = 1.458 \times 10^{-6} \frac{\text{N s}}{\text{m}^2}$, $C_2 = 110.4$ K, and R is the specific gas constant which is equal to 287.05 $\frac{\text{J}}{\text{kg K}}$. The reference temperature is also used to determine the constant pressure specific heat capacity using the following semi-empirical formula [18]:

$$c_p = C_0 + C_1 T^* + C_2 T^{*2} + C_3 T^{*3} + C_4 T^{*4} \tag{14}$$

where the constants are given in Table 2.

Table 2: Equation (14) Constants

C_0	1.03409
C_1	-0.284887×10^{-3}
C_2	0.7816818×10^{-6}
C_3	$-0.4970786 \times 10^{-9}$
C_4	$0.1077024 \times 10^{-12}$

After the reference properties are calculated, the Reynolds number at the reference state is calculated:

$$Re_x^* = \frac{\rho^* V_e x_d}{\mu^*} \tag{15}$$

where V_e is the air velocity magnitude at the boundary layer edge and x_d is the downstream distance from the leading edge. Next, the reference skin friction coefficient, C_f^* is determined. C_f^* is highly dependent on the flow regime.

$$C_f^* = \frac{0.664}{\sqrt{Re_x^*}} \quad \text{(laminar)} \tag{16}$$

$$C_f^* = \frac{0.0592}{(Re_x^*)^{1/5}} \quad (\text{turbulent}, 5 \times 10^5 \leq \text{Re}_\text{x}^* \leq 10^7) \tag{17}$$

$$C_f^* = \frac{0.370}{(\log_{10} Re_x^*)^{2.584}} \quad (\text{turbulent}, 10^7 \leq \text{Re}_\text{x}^* \leq 10^9) \tag{18}$$

Because the flow is assumed to be turbulent, only Equations (17) and (18) are used to calculate the skin friction coefficient. Next, the Stanton number is calculated:

$$St^* = \frac{C_f^*}{2} Pr^{-2/3} \tag{19}$$

The Stanton number describes the coupling of the air friction and the convective heat transfer. Lastly, the heat flux coefficient is found using,

$$h = St^* C_p \rho^* V_e \tag{20}$$

The heat flux coefficient is used in Equation (6) to calculate the surface heat flux.

3.4 Kriging Surrogates

Kriging, also known as the spacial Best Linear Unbiased Prediction (BLUP) [19], is an interpolation method that uses data points at pre-known locations within a defined design space and builds an optimized, interpolative model that allows for predictions between data points [20][19]. This tool is useful in applications where data points are expensive or difficult to obtain, but an estimation of an unknown, underlying function is needed [21]. The benefit of using kriging surrogates is that they can be orders of magnitude faster than the expensive primary source. The trade-off is that there is a decrease of accuracy when predicting away from the known data points. Kriging relies on some important assumptions:

1. Kriging surrogates are continuous [21].
2. Most types of kriging treats the objective function as deterministic [21]. In other words, kriging does not account for noise, biases, and other errors from the primary data source. Note that if the output data is nondeterministic, it can be modeled using nondeterministic kriging.
3. The spacial field (design space) is a realization of a random, stochastic process [20]. This assumes the *locations* of the data points are considered random.

For example, one objective in this research is to describe the lift coefficient of a hypersonic vehicle using CFD data. The design space is three-dimensional where the spacial variables are the Mach number, angle of attack, and flight altitude of the airflow. The objective function of interest can be described as $C_L(M, \alpha, h)$. CFD provides solutions for $C_L(M, \alpha, h)$, but only at a finite number of discrete points. Kriging has the capability of building a continuous model using multiple $\mathrm{C_L}$ values throughout the design space. This allows for the estimation of the continuous $C_L(M, \alpha, h)$ function while using discrete training data points.

The kriging surrogate model can be roughly generalized using the form,

$$\hat{y}(\mathbf{x}) = f(\mathbf{x}) + Z(\mathbf{x}) \tag{21}$$

where, $\hat{y}$ is the prediction value, $\mathbf{x}$ is a design point with a size of k design variables (dimensions), $f(\mathbf{x})$ is a baseline function, and $Z(\mathbf{x})$ is a Gaussian process with a mean of 0 and a varience of σ^2 [22]. This shows that kriging is completed in two steps: first, a regression function or constant value is computed using the data to get $f(\mathbf{x})$, and secondly, a stochastic, Gaussian function is built using the residuals of the data points for $Z(\mathbf{x})$. *Ordinary kriging* assumes that the regression function, $f(\mathbf{x})$ is a constant value calculated from the mean of the data while *universal kriging* sets $f(\mathbf{x})$ as a regression function with a form,

$$f(\mathbf{x}) = \sum_{i=1}^{p} \alpha_i b_i(\mathbf{x}) \tag{22}$$

where, α_i are the polynomial coefficients, and b_i as the $i = 1 \dots p$ order polynomial terms.

3.4.1 History of Kriging

The method of kriging was coined after the originator of the method, Danie G. Krige. Krige was involved in surveys of gold fields and gold mine valuations in South Africa in the 1940's and 50's. Previously, gold mine locations were chosen from rudimentary grade estimates that used "simple arithmetic and empirical formulae guided by practical experience" [23]. Typically, several core-samples of ore throughout a region would be collected and processed and the sample mean of the gold grade was taken to estimate the overall grade in that region. Krige observed that by using this approach there were large disparities between estimated grade and actual grade within regions and this increased the risk of selecting a gold mine location in an area of low-grade ore, hurting profits. Krige understood that the local distribution of ore grade in a region needed to be evaluated in order to mine high-grade locations consistently. This led Krige to develop a statistical method to locally predict high-grade and low-grade ore regions which he introduced in his masters thesis in 1951 [19].

Krige continued to develop his statistical methodology throughout the 1950's exclusively for the application of gold mine valuation. In 1963, the French mathematician Georges Matheron formally developed Krige's statistical method in his *Principles of Geostatistics* [24]. Matheron expanded the scope of the theory by including the spacial component [19]. The comprehensive formalization of kriging done by Matheron opened the door for additional applications of the theory. In 1989, Jerome Sacks further extended the application of kriging to computer simulation in *Design and Analysis of Computer Experiments*. Kriging applies very well to computer simulations because they are inherently deterministic, or very nearly deterministic.

3.4.2 Kriging Theory

The following discussion will follow the development of kriging found in *Engineering Design via Surrogate Modelling* [21] and the ooDACE Toolbox users manual [22].

Kriging models are built around correlation functions, also known as radial basis functions or kernels. There are many types of correlation functions, but they each perform a common utility. Correlation functions calculate the statistical correlation between two random points and are solely dependent on the spacial distance between those two points [22]. The closer two points are together, the more correlated they are. Two points that are highly correlated with each other are more influenced by each other's prediction values. Correlation functions, ψ, must meet the following criteria [21]:

1. ψ is continuous
2. ψ is symmetric, $\psi(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) = \psi(\mathbf{x}^{(2)}, \mathbf{x}^{(1)})$
3. ψ is positive definite

The most common form of correlation function in kriging is defined as,

$$\psi(\mathbf{x}^{(i)}, \mathbf{x}^{(l)}) = \exp\left(-\sum_{j=1}^{k} \theta_j \left| x_j^{(i)} - x_j^{(l)} \right|^{p_j}\right) \tag{23}$$

where k is the number of design variables (length of $\mathbf{x}$), θ_j is the width parameter, p_j is the order parameter, and i and l are indices of the two points being measured. The order parameter, p_j, determines the initial drop in correlation as distance increases. p_j specifies the "pointedness" or "smoothness" of the correlation around the distance equals 0 point. A small order (say, $p = 0.5$) initially decreases the correlation between two points quickly as distance increases, while a larger order (say, $p = 2$) initially decreases the correlation between two points slowly as distance increases. Some examples of one-dimensional correlation functions with varying order parameters are demonstrated in Figure 5. Typically, the order parameter is kept constant across the entire design point space.

The set of width parameters, θ_1,...,θ_k, also called the *hyperparameters*, describes the area of influence about a point on nearby points in a similar dimension. The hyperparameters define how quickly the correlation drops between two points as the distance increases for a given dimension. Large θ_i values correspond to low correlation between points, typically indicative of nonlinearity in the ith dimension. Whereas a smaller θ_i means that a given point is correlated more closely with other points further away, indicating linearity in the ith dimension. A demonstration of the effect of θ on the correlation function is shown in Figure 6. The hyperparameters, θ_1,...,θ_k, are calculated using Maximum Likelihood Estimation (MLE), an optimization problem that finds optimal values for θ. MLE will be discussed later.

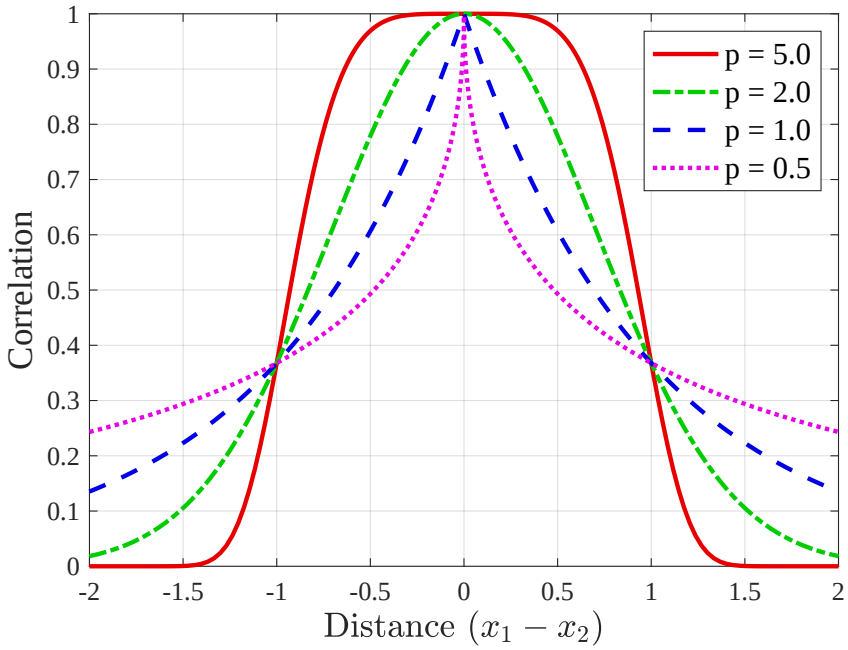


Figure 5: One-dimensional correlation function examples with varying order parameters, $\theta = 1$.

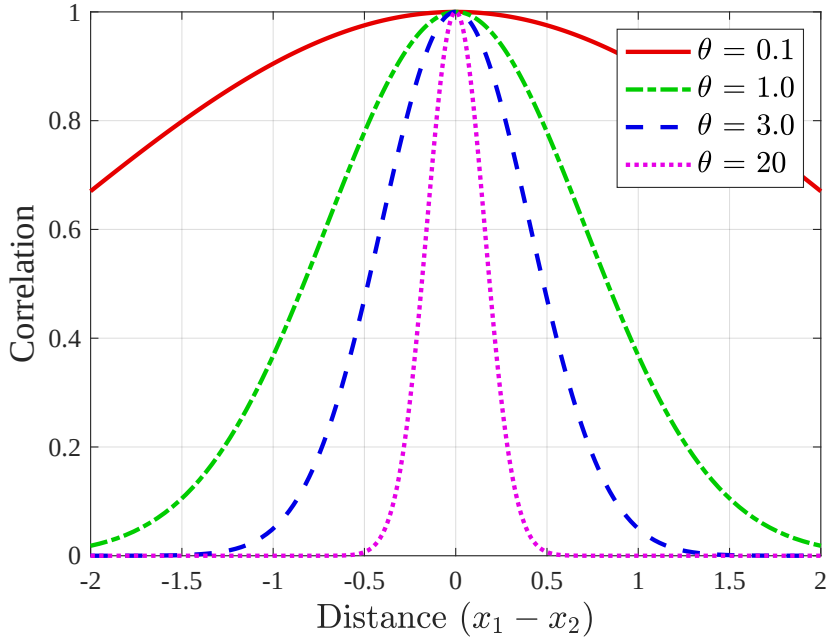


Figure 6: One-dimensional correlation function examples with varying width parameters, $p = 2$.

As mentioned before, there are many types of correlation functions that serve the same purpose, but use slightly different mechanisms. Another common correlation function structure is the

Matern function. The Matern function has the general form,

$$\phi_\nu(\mathbf{x}^{(i)}, \mathbf{x}^{(l)}) = (r)^\nu \mathcal{K}_\nu(r) \tag{24}$$

where, $r = \sqrt{\sum_{j=1}^{k} \theta_j (x_j^{(i)} - x_j^{(l)})^2}$, and $\mathcal{K}_\nu$ is a modified Bessel function with an order of ν. Two common ν constants are $\frac{3}{2}$ and $\frac{5}{2}$ which correspond to the Matern-32 and Matern-52 functions [22]:

$$\phi_{\nu=3/2}(\mathbf{x}^{(i)}, \mathbf{x}^{(l)}) = \left(1 + \sqrt{3}r\right) \exp\left(-\sqrt{3}r\right) \tag{25}$$

$$\phi_{\nu=5/2}(\mathbf{x}^{(i)}, \mathbf{x}^{(l)}) = \left(1 + \sqrt{5}r + \frac{5r^2}{3}\right) \exp\left(-\sqrt{5}r\right) \tag{26}$$

Some other examples of different correlation functions are spline, spherical, linear, cubic, and multiquadric. A variety of correlation functions will be evaluated and compared in Chapter 5.

To build a kriging model, a set of spacial sample points, $\mathbf{X}$, are defined along with their corresponding responses, $\mathbf{y}$:

$$\mathbf{X} = \begin{Bmatrix} \mathbf{x}^{(1)} \\ \vdots \\ \mathbf{x}^{(n)} \end{Bmatrix} \qquad \mathbf{y} = \begin{Bmatrix} y^{(1)} \\ \vdots \\ y^{(n)} \end{Bmatrix}$$

where n is the number of training points. In the scope of this research, $\mathbf{X}$ describes the design space of Mach number, angle of attack, and altitude, while $\mathbf{y}$ describes the aerodynamic responses calculated from the CFD simulation. The observed responses will be treated as originating from a stochastic process, despite the responses most likely originating from a deterministic process in reality [21]. These "stochastic" responses are denoted in the set of random vectors, $\mathbf{Y}$.

$$\mathbf{Y} = \begin{Bmatrix} Y(\mathbf{x}^{(1)}) \\ \vdots \\ Y(\mathbf{x}^{(n)}) \end{Bmatrix}$$

A correlation function is selected and the correlation matrix $\mathbf{\Psi}$ is calculated:

$$\text{cor}\left[Y(\mathbf{x}^{(i)}), Y(\mathbf{x}^{(l)})\right] = \psi = \exp\left(-\sum_{j=1}^{k} \theta_j \left|x_j^{(i)} - x_j^{(l)}\right|^{p_j}\right) \tag{27}$$

$$\mathbf{\Psi} = \begin{bmatrix} \text{cor}\left[Y(\mathbf{x}^{(1)}), Y(\mathbf{x}^{(1)})\right] & \dots & \text{cor}\left[Y(\mathbf{x}^{(1)}), Y(\mathbf{x}^{(n)})\right] \\ \vdots & \ddots & \vdots \\ \text{cor}\left[Y(\mathbf{x}^{(n)}), Y(\mathbf{x}^{(1)})\right] & \dots & \text{cor}\left[Y(\mathbf{x}^{(n)}), Y(\mathbf{x}^{(n)})\right] \end{bmatrix} \tag{28}$$

The next step is to select the set of hyperparameters that maximize the prediction likelihood using the MLE process. This portion is the most computationally expensive portion of kriging. The likelihood function, L, is,

$$L(\mathbf{Y}^{(1)}, \dots, \mathbf{Y}^{(n)}|\mu, \sigma) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{\sum(\mathbf{Y}^{(i)} - \mu)^2}{2\sigma^2}\right] \tag{29}$$

where $\mathbf{Y}$ is the set of observations, μ is the observation mean, and σ^2 is the observation variance. The likelihood function measures the goodness of fit of a statistical model to the sample data. The function can be rewritten in terms of the sample data:

$$L = \frac{1}{(2\pi\sigma^2)^{n/2}|\mathbf{\Psi}|^{1/2}} \exp\left[-\frac{(\mathbf{y} - \mathbf{1}\mu)^T\mathbf{\Psi}^{-1}(\mathbf{y} - \mathbf{1}\mu)}{2\sigma^2}\right] \tag{30}$$

where $\mathbf{1}\mu$ is the mean of $\mathbf{Y}$. This expression can be simplified by taking the natural log of both sides,

$$\ln(L) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2}\ln|\mathbf{\Psi}| - \frac{(\mathbf{y} - \mathbf{1}\mu)^T\mathbf{\Psi}^{-1}(\mathbf{y} - \mathbf{1}\mu)}{2\sigma^2} \tag{31}$$

The MLEs are computed by taking the derivative of Equation (31) and setting it equal to 0. The MLEs, μ and σ^2, to maximize the likelihood are,

$$\hat{\mu} = \frac{\mathbf{1}^T\mathbf{\Psi}^{-1}\mathbf{y}}{\mathbf{1}^T\mathbf{\Psi}^{-1}\mathbf{1}} \tag{32}$$

$$\hat{\sigma}^2 = \frac{(\mathbf{y} - \mathbf{1}\mu)^T\mathbf{\Psi}^{-1}(\mathbf{y} - \mathbf{1}\mu)}{n} \tag{33}$$

These MLEs are substituted back into Equation (31) and the constant terms are removed to yield the *concentrated ln-likelihood function*,

$$\ln(L) \approx -\frac{n}{2}\ln(\hat{\sigma}^2) - \frac{1}{2}\ln|\boldsymbol{\Psi}| \tag{34}$$

The first term, $-\frac{n}{2}\ln(\hat{\sigma}^2)$ represents the goodness of fit while the second term, $-\frac{1}{2}\ln(\boldsymbol{\Psi})$ represents the complexity of the model. The concentrated ln-likelihood function is dependent on the unknown hyperparameters, $\boldsymbol{\theta}$ (and $\mathbf{p}$ if its not specified). Since we want to maximize the likelihood, we set up the following optimization problem:

$$\max_{\boldsymbol{\theta},\mathbf{p}} \quad -\frac{n}{2}\ln(\hat{\sigma}^2) - \frac{1}{2}\ln|\boldsymbol{\Psi}| \tag{35}$$

This problem is very difficult to solve because the function is not directly differentiable with respect to $\boldsymbol{\theta}$ and $\mathbf{p}$. Because of this, a numerical optimization scheme is needed to maximize $\ln(L)$. Typically, a global optimization scheme is used to solve for the hyperparameters. After accomplishing MLE, we now have solutions for $\hat{\sigma}^2$, $\hat{\mu}$, $\hat{\boldsymbol{\theta}}$, and $\hat{\mathbf{p}}$ that maximize the likelihood, and now the kriging prediction function can be built.

Say a new prediction, $\hat{y}$, is needed at $\mathbf{x}$. This prediction must be consistent with the existing sample data and the correlations determined earlier. Thus, the prediction needs to maximize the likelihood of the sample data and the prediction itself. This is done by creating an augmented vector with the sample data and the prediction: $\tilde{\mathbf{y}} = \{\mathbf{y}^T, \hat{y}\}^T$. A vector of correlations between the prediction location and the sample data is created,

$$\boldsymbol{\psi} = \begin{bmatrix} \mathrm{cor}[Y(\mathbf{x}^{(1)}), Y(\mathbf{x})] \\ \vdots \\ \mathrm{cor}[Y(\mathbf{x}^{(n)}), Y(\mathbf{x})] \end{bmatrix} = \begin{bmatrix} \psi^{(1)} \\ \vdots \\ \psi^{(n)} \end{bmatrix} \tag{36}$$

An augmented correlation matrix is now defined:

$$\tilde{\boldsymbol{\Psi}} = \begin{bmatrix} \boldsymbol{\Psi} & \boldsymbol{\psi} \\ \boldsymbol{\psi}^T & 1 \end{bmatrix} \tag{37}$$

Another ln-likelihood function is generated for the augmented matrix,

$$\ln(L) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\hat{\sigma}^2) - \frac{1}{2}\ln|\tilde{\boldsymbol{\Psi}}| - \frac{(\tilde{\mathbf{y}} - \mathbf{1}\hat{\mu})^T \tilde{\boldsymbol{\Psi}}^{-1} (\tilde{\mathbf{y}} - \mathbf{1}\hat{\mu})}{2\hat{\sigma}^2} \tag{38}$$

Only the last term is dependent on $\hat{y}$, so this will be the only term considered in the likelihood maximization. Removing the unneeded terms and substituting in expressions for $\tilde{\mathbf{y}}$ and $\tilde{\boldsymbol{\Psi}}$ yields:

$$\ln(L) \approx \frac{-\begin{bmatrix}\mathbf{y}-\mathbf{1}\hat{\mu}\\ \hat{y}-\hat{\mu}\end{bmatrix}^T \begin{bmatrix}\boldsymbol{\Psi} & \boldsymbol{\psi}\\ \boldsymbol{\psi}^T & 1\end{bmatrix}^{-1} \begin{bmatrix}\mathbf{y}-\mathbf{1}\hat{\mu}\\ \hat{y}-\hat{\mu}\end{bmatrix}}{2\hat{\sigma}^2} \tag{39}$$

This expression is maximized with respect to the prediction, $\hat{y}$. The final MLE for $\hat{y}$ is shown in Equation (40). Reference Forrester [21] for a thorough mathematical explanation of the maximization.

$$\boxed{\hat{y}(\mathbf{x}) = \hat{\mu} + \boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1}(\mathbf{y} - \mathbf{1}\hat{\mu})} \tag{40}$$

This equation is called the *kriging equation*, and it allows for interpolation between spatially sparse data. After the kriging equation is computed using MLE, making predictions using this equation is computationally marginal.

It is important to note that there are some potential issues with kriging that must be understood before implementation. First, the correlation matrix must be invertible which means the matrix must be positive definite. This typically occurs if there are linear dependencies between subsets of data points and/or there exists very close spacing between data points. This can be mitigated by taking the pseudo-inverse of the correlation rather than the direct inverse and by sampling randomly rather than uniformly. Secondly, kriging does not perform well with large quantities of training points. Model errors can build up with very large numbers of points and the correlation matrix may contain negative eigenvectors. The spacial domain can be split into different portions and separately modeled to help this issue. Finally, kriging is sensitive to dimensionality, and it becomes computationally infeasible with many design variables. A solution to this issue is using design variable linking.

3.4.3 Kriging Application in Aerodynamics

There were several notable studies conducted that utilized kriging to model the aerodynamics of a vehicle. The paper that was primarily referenced for the research in this research was *Rapid Steady-State Hypersonic Aerothermodynamic Loads Prediction Using Reduced Fidelity Models* by Dreyer *et.al.* [25].

3.5 Sampling Plans

Before collecting the training data for the surrogate, a sampling plan, or Design of Experiments (DoE), needed to be selected. A uniform level of accuracy is desired within the design space, and this is accomplished by a uniform spread of points [21]. An optimal, space-filling DoE needs to satisfy two criteria: first, the points must be uniformly distributed throughout the design space (low discrepancy); and second, it needs to exhibit stratification which refers to the uniformity of projections of the data points on each design variable axis [26]. Additionally, it is preferred that

the set of points be ordered in a way that minimizes the discrepancy no matter the number of sample points. There are several options of DoE types. Some common examples are listed below and each example is illustrated in Figure 7 for two design variables, X_1 and X_2. Note, the first 8 points in each sequence is colored red.

1. *Full Factorial* - This sampling plan places points on the design space boundaries and evenly fills the interior in a linear, structured manner. This DoE only operates if the number of design points is a perfect k-dimensional root, *i.e.* $N = x^k$ with x being a whole number. The full factorial excels in distributing points uniformly (the discrepancy is minimal), but it fails in stratification because the projections of the points on either axis overlap each other [21].

2. *Latin Hypercube* - Latin hypercubes optimize the stratification of the design space [21]. A latin hypercube algorithm divides the design space into a number of equally sized bins (hypercubes) and points are placed inside the bins such that there are no other bins parallel that contain points. This method of placing points ensures that the projections of the points on any axis never overlap. While the entire sequence satisfies the uniformity and stratification criteria, the order of the points in the sequence is random, and a subset of a latin hypercube sequence is not guaranteed to be uniform.

3. *Sobol Set* - This strategy uses a quasi-random base-2 digital process to fill the design space uniformly. It operates similarly to the latin hypercube in that the design space is split into 2^N bins and points are allocated into the bins such that the stratification is upheld. The Sobol set performs better than the latin hypercube in that the discrepancy of the set is small regardless of the subset selected. In other words, the first 10, 20, 50, 100, etc. points of a common Sobel set each fills the design space uniformly. In Figure 7, the red points for the Sobol set show the first 8 points in the sequence, and they are generally placed so that they are space-filling. This is also true of all 36 points (red+black). The algorithm used to generate a Sobol set is described in detail in *ALGORITHM 659 Implementing Sobol's Quasirandom Sequence Generator* [27].

4. *Random Process* - This process uses a generic random number generator to determine the locations of points. This is not a preferred method because the uniformity and stratification of points is not guaranteed.

3.6 The Generic Hypersonic Vehicle

3.6.1 GHV Initial Design

In 2012, Air Force Research Laboratory (AFRL) researchers Brent Ruttle, Jacob Stork, and Glenn Liston wrote a technical report about conceptual hypersonic vehicle design [28]. They stated that in the field of hypersonic vehicle design outside of AFRL there is a distinct lack of accessibility to these designs due to ITAR or proprietary restrictions. Additionally, the tools and models used by AFRL researchers were also unavailable to academia and outside research. This led them to develop a vehicle design using readily accessible tools with the purpose of the design being

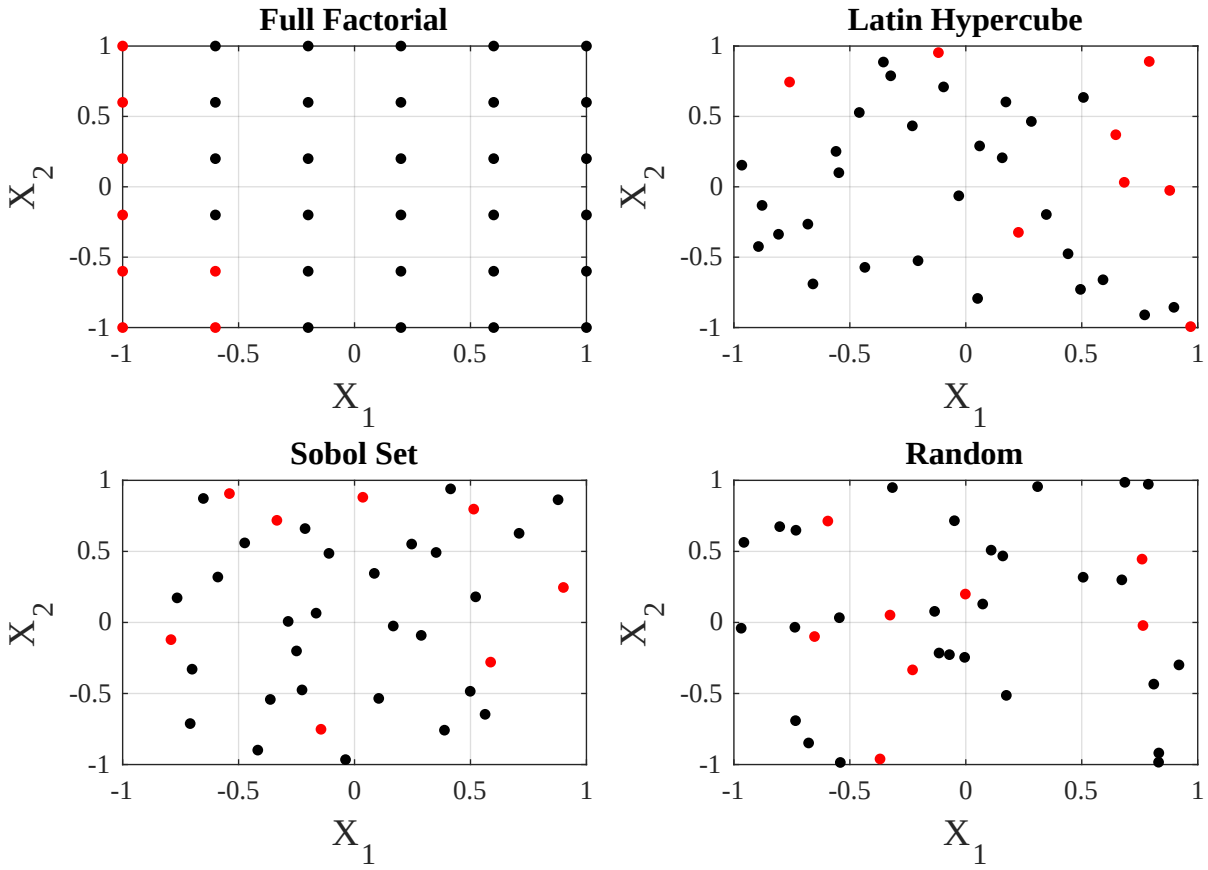


Figure 7: Types of Sampling Plans for 36 Points

publicly releasable and relevant to current hypersonics research. Because of the looser restrictions on this design, academic or industry partners could collaborate more in the field of hypersonic vehicle design and thus expand the knowledge of the field outside of restriction.

Ruttle et al. developed the Generic Hypersonic Vehicle (GHV) design based on the overall need stated above. The vehicle was developed considering the tight integration of aerothermodynamic, structural, controllability, operability, and propulsion systems common in all hypersonic vehicles. This resulted in a design featuring a blended wing body built around a scramjet engine. The vehicle design methodology contained three main phases:

1. *Initial design study.* This phase defined the design cruise conditions, the mission profile, list of desired maneuvers, and others. The range, weight, lift, etc. for the vehicle was estimated using simple relationships and the predefined conditions. A summary of the initial design parameters for the GHV is listed in Table 3. The vehicle was to be designed with the capabilities of Mach 6 cruise at dynamic pressures ranging from 1000 to 2000 psf, sustained 2 g load factor, and a range of 1200 nmi.

Table 3: GHV Initial Design Study Parameters

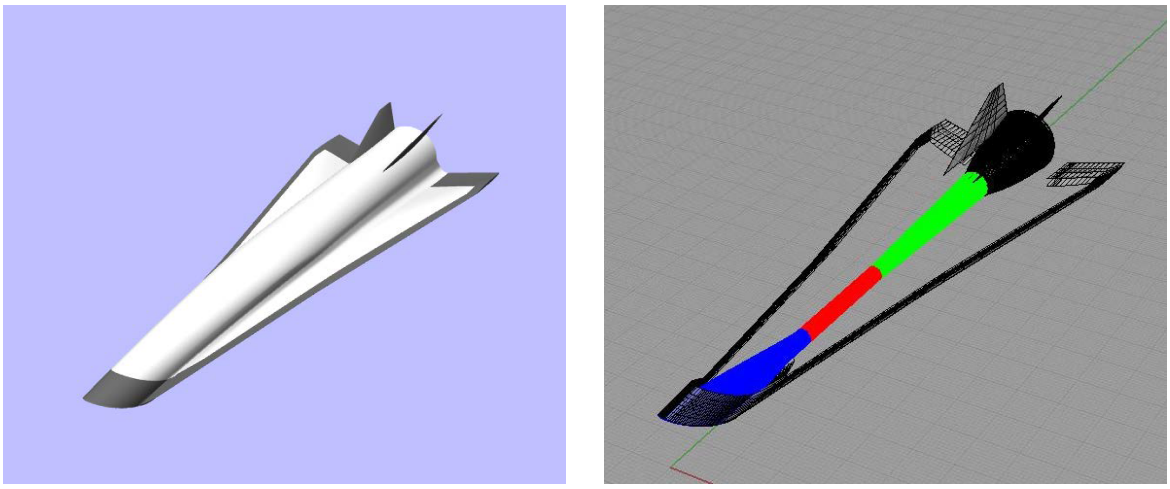
Cruise Mach number	6
Cruise dynamic pressure (psf)	1500
Max lift to drag	5
Cruise specific impulse (s)	1000
Fuel mass fraction	0.39
Weight (lbf)	5000
Thrust (lbf)	2000
Fuel mass flow rate (lbm/s)	2
Air mass flow rate (lmb/s)	30

2. *Parametric engine design.* The scramjet engine was designed using the parameters defined in the initial design study. Experimental scramjet data was used to generate engine performance maps across internal Mach numbers ranging from 3 to 5 and AoA ranging from -4° to 6°. The estimated GHV engine performance and dimensions were found using the initial design parameters and the experimental performance maps.
3. *Vehicle Design.* Using the vehicle and engine performance determined prior, five unique vehicle concepts were developed for different flowpath scales ranging from 1X to 5X (10 to 50 lbm/s). The GoHypersonic ECO software suite was used to develop the vehicle concepts. The design performance was calculated and compared to the expected performance from the first two phases and the concept was redesigned if necessary.

Figure 8 shows an external CAD model and the internal configuration of the Road Runner GHV concept resulting from the methodology discussed above. The vehicle consists of two vertical stabilizers and a highly swept wing. The entire blended wing geometry conforms to the shape of the scramjet engine. The scramjet engine consists of four main components: the inlet (blue), isolator (red), combustion chamber (green), and nozzle (black). These scramjet components span across the entire length of the vehicle.

3.6.2 GHV Parameterization and Previous Work

The GHV concept was adopted by others in the years following the initial design analysis as a test subject for many novel analyses. Ferguson et al. from North Carolina A&T State University studied the aerodynamics of the GHV scramjet engine in detail in 2015 [29]. They used inviscid, laminar, and turbulent methods using the AVUS and US3D CFD codes to inspect the aerodynamic behavior of the scramjet engine. The shock train inside the isolator, the fuel injector in the combustor, and overall profile of the scramjet were implemented in their CFD model and analyzed. Later, in 2018, Ferguson et al. studied the aerothermodynamic performance of the



(a) Top Down View (b) Internal Configuration

Figure 8: Road Runner GHV Concept [28]

entire GHV by modeling the internal and external configuration [30]. They developed simple profiles of lift and drag with respect to angle of attack using similar CFD methods as before.

The GHV concept was later used by Sagerman et al. from University of Dayton and AFRL as a test subject to assess the aerothermodynamic effects on the vehicle [31]. They ran a comparison study on the GHV and two other complex hypersonic vehicle configurations each at Mach 6 flight conditions. The study used experimental data using the Mach 6 tunnel at Wright-Patterson AFB and CFD data for comparison. The thermal behavior of the wall was inspected using a temperature sensitive paint coating the experimental models. The experimental GHV configuration from Sagerman is shown in Figure 9 and the surface temperature contours from the thermal paint are shown in Figure 10. Note, the highest temperatures were located around the edge of the inlet and the wing/tail leading edges. The CFD analysis used panel method, Euler, Euler boundary-layer, and Navier-Stokes codes to predict the aerodynamic behavior of the GHV at different fidelities. Unfortunately, very little quantitative data for the GHV that could be used for comparison was published in the paper.

In a study conducted by James Haley and others at the University of Texas in 2018, the sizing of the GHV design developed by Ruttle, Stork, and Liston was verified [32]. The GHV was analyzed using a system called the Database Management System that generated a sizing code for the vehicle given a mission profile and performance metrics. The mission was described in greater detail to include an air-launch, booster stage ignition, maneuvering, and vehicle recovery. The main purpose of this work was to evaluate the utility of the software, but they determined that the GHV concept was feasible for the mission profile, cruise conditions, and sizing.

Later work conducted by Ian Boyd, Ramana Grandhi, and others from AFIT and AFRL focused

Figure 9: GHV Wind Tunnel Model with Thermal Sensitive Paint, from Sagerman et al. [31]

on the optimization of the geometry and structure of the GHV [33]. The team converted the GHV geometry into a fully parameterized configuration for the software to accomplish the optimization. The parameterization was accomplished using Engineering Sketch Pad (ESP) which generated high-quality CAD models to be used for CFD and Finite Element Analysis (FEA). The parameterization fully defined the geometric dimensions of the fuselage, wings, vertical tail, and scramjet using 50 unique variables. The aerodynamic loads on the vehicle were predicted using the CBAero panel method code and the structural optimization was conducted using ASTROS. The optimization problem sought to minimize the mass of the GHV structure while considering all of the aerodynamic and gravitational loads on the vehicle. The result from this research was a structurally optimized GHV geometry which is shown in Figure 11. This optimized GHV configuration is used as the CFD test subject in this research.

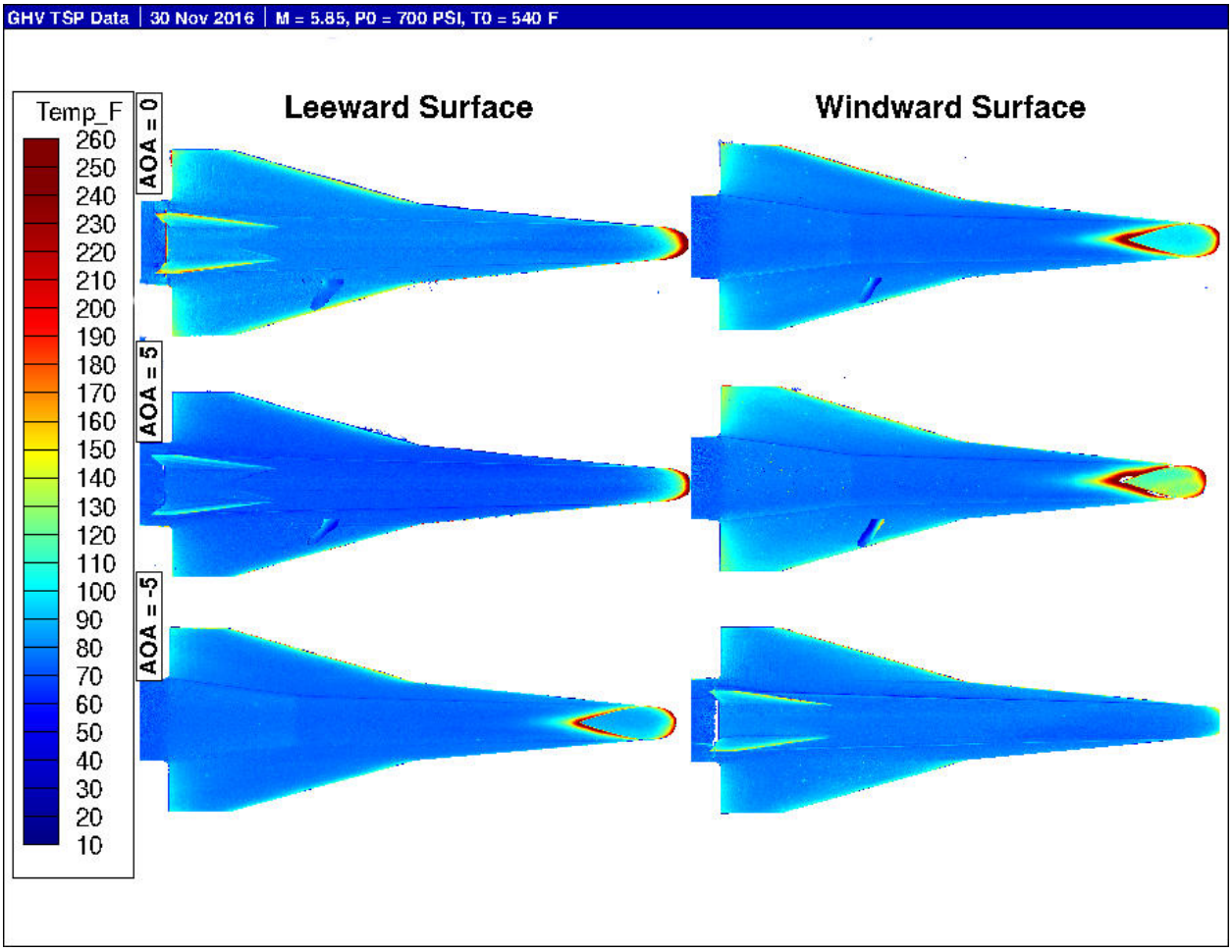
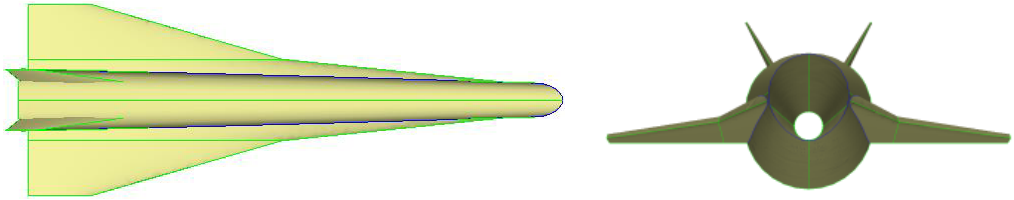


Figure 10: GHV Wind Tunnel Thermal Contours [31]



(a) Top (b) Front

Figure 11: Optimized GHV Configuration [33]

4 METHODOLOGY

4.1 Overview

The main objective for this research was to use surrogate modeling of CFD data from the GHV in order to rapidly evaluate its aero-thermal characteristics in hypersonic flight. This was accomplished for the purpose of predicting vehicle performance within the preliminary/conceptual vehicle design phase. The CFD data was collected using the Euler equation set which modeled inviscid effects such as shocks, isentropic compression, flow expansion, etc. Euler did not model viscous features such as boundary layers and turbulence. Euler CFD was assumed to be the highest justifiable fidelity for use within the preliminary/conceptual design phase.

As described in Chapter 3, the GHV was conceived by the research conducted by Ruttle, Stork, and Liston [28]. The concept was created because of the need for a publicly releasable hypersonic vehicle design to spur collaboration in the field and to better develop preliminary design methodology for hypersonic vehicles. The concept has the purpose of being a testbed for conceptual hypersonic vehicle design using multi-fidelity and multi-disciplinary tools. Ian Boyd, Corey Fischer, and Ramana Grandhi took the GHV geometry from Ruttle et al. and used a novel design framework which optimized the GHV geometry using pressure loading configurations [33]. The optimized GHV geometry from this research was used as the test subject for the surrogate modeling methodology in this research.

The overall process involved the following steps:

1. Define design variables and their boundaries.
2. Generate a space-filling set of sample points within the design space.
3. Run CFD simulations at each of the sample points and their conditions.
4. Build kriging surrogates of CFD output parameters.
5. Assess the suitability of surrogate models using validation data points.

This chapter will discuss the detailed methodology behind each of these steps and provide reasoning behind the decisions made. The discussion will start with the CFD process which includes defining the geometry, generating a grid, and computing the solution. The discussion will then move on to how the sample space was defined and how the kriging process was conducted.

4.2 Geometry

The GHV geometry used for this study is shown in detail in Figure 12 along with primary dimensions. The wing is comprised of a strake with a 84° sweep angle and a main wing with a 73° sweep angle. A sample cross section of the wing is shown in red. The wings feature a slightly cambered biconvex cross section and the vertical tails have a symmetrical biconvex cross section . A detailed list of the GHV dimensions is tabulated in Table 4. Figure 13 shows an isometric view of the GHV. A side 2-D projection of the internal scramjet flow path is depicted in Figure 14

which highlights the scramjet components and dimensions. The isolator, combustor, and nozzle each have circular cross sections.

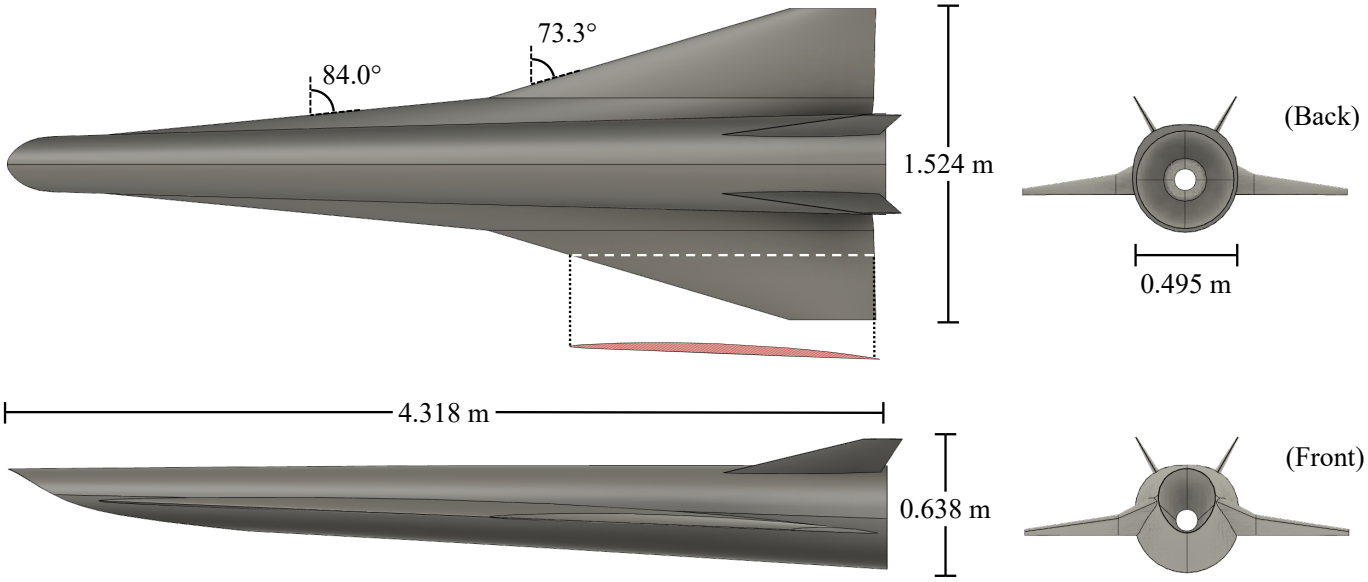


Figure 12: GHV Geometry

4.3 Grid Generation

4.3.1 Surface Grid

The surface grid of the GHV was chosen to be structured so that the curvature of the body could be accurately defined and so the cell spacing could be kept more uniform. The surface grid only modeled half of the vehicle cut at the symmetry plane because the sideslip angle was to be held at 0° for each CFD simulation. This reduced the required number of grid cells by half which reduces computational time while not affecting the flow solution. Link3D was used to generate the surface grid because of its excellent capabilities in creating quality fully structured grids for complex geometries.

The surface grid was defined by a complex topology containing 195 rectangular panels projected to the GHV curvature (fig. 15). The topology structure was made manually in Link3D in a configuration that distributed cells as evenly as possible across the surface. Each panel side was assigned a number of cells dependent on the spacing required and the region. Note, since the surface grid was structured, each panel was required to contain equal number of cells on opposing sides so a uniform rectangular arrangement of cells could be defined. After the topology was defined and all sides assigned a cell number, the grid was smoothed by an elliptic Laplace PDE

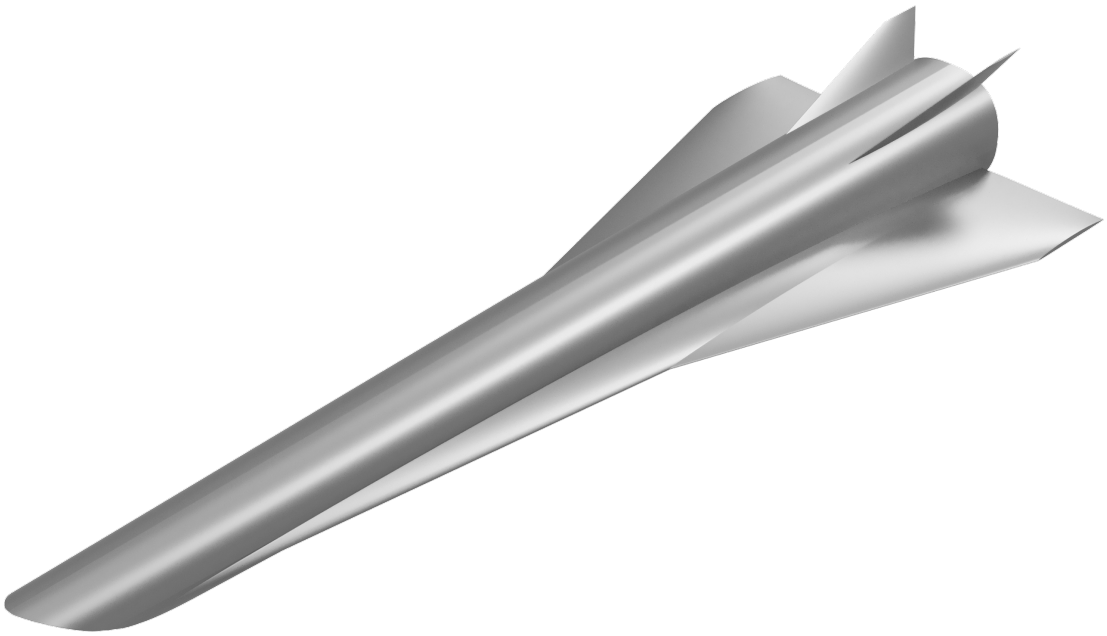


Figure 13: GHV Isometric Rendering

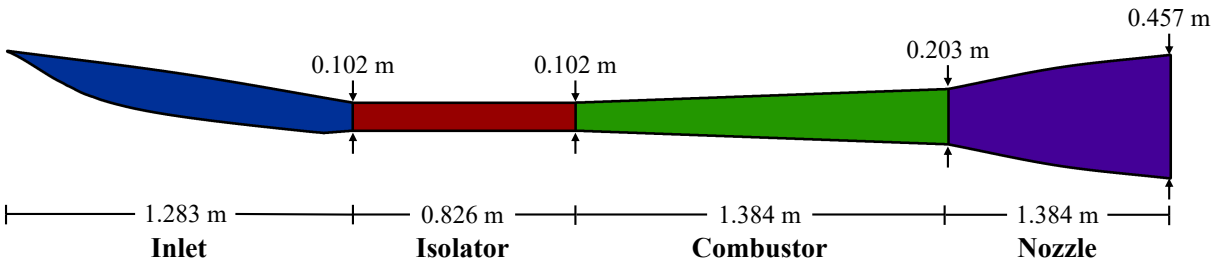


Figure 14: GHV Scramjet Geometry

solver inside Link3D. The smoother moved the grid cells along the surface to evenly distribute the cells and enforce continuous spacing.

The process of choosing the topology and assigning cell numbers to each edge was extremely demanding due to the qualitative judgement needed and the complex geometry. The topology needed to wrap around each surface feature including the tail, wing, and internal configuration. The Laplace PDE solver was extremely sensitive to the cell spacings and topology input given by the user. It was common for the iterative solver to diverge and crash while smoothing the grid. When the grid diverged during the smoothing process, it occured as a result of the specified grid spacings being too large. Link3D performed better when generating finer girds. The finalized surface grid was obtained after many attempts, and an example is shown in Figure 16. Note that

Table 4: GHV Principal Dimensions

Fuselage Length (m)	4.318
Overall Length (m)	4.392
Fuselage diameter, nose (m)	0.270
Fuselage diameter, tail (m)	0.495
Vehicle Planform Area (S, m^2)	3.275
Wingspan (m)	1.524
Strake Root Chord (m)	3.807
Mid-Wing Chord (m)	1.903
Wing Tip Chord (m)	0.431
Strake LE Sweep Angle (°)	84.0
Wing LE Sweep Angle (°)	73.3
Wing TE Sweep Angle (°)	0.0
Average Wing Incidence (°)	2.39
Tail Length (m)	0.186
Tail Root Chord (m)	0.766
Tail Tip Chord (m)	0.189
Tail LE Sweep Angle (°)	74.8
Tail TE Sweep Angle (°)	29.6
Tail Tilt Angle (°)	29.9
Inlet Capture Area (m^2)	0.0624
Combustor Expansion Angle (°)	2.1
Nozzle Expansion Angle (avg, °)	8.7

the grid in the figure was coarsened so the viewer could see the cells clearly. The figure is not representative of the true cell spacings due to the coarsening. Figures A.1(a) to A.1(d) show non coarsened views of the surface grid which is difficult to properly display.

After the surface grid was generated, additional grids of varying resolutions were solved for to be used in the grid convergence study. An original surface grid was created using the largest spacing allowed for the grid to converge. Link3D contained a convenient feature that uniformly changes the cell spacing across the grid by a user defined factor. A series of finer surface grids containing different cell counts were subsequently generated using this method. This set of surface grids was used later in the grid convergence study in Section 4.4.

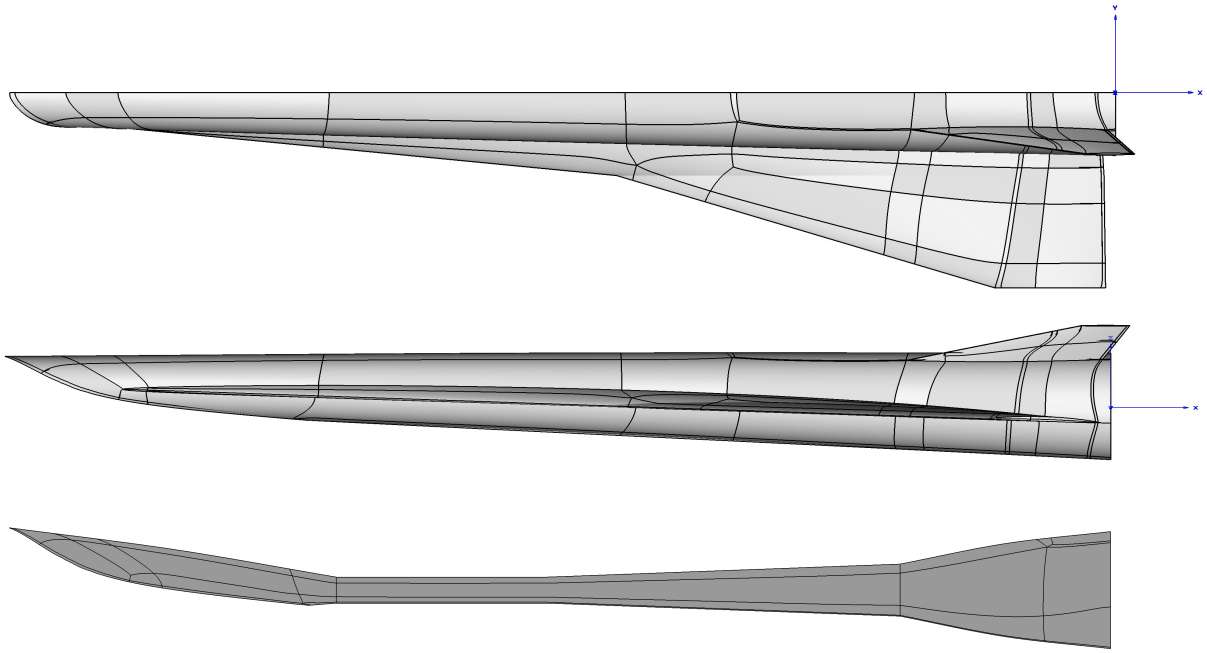


Figure 15: Surface Grid Topology

4.3.2 Volume Grid

During the development of the volumetric grid for the GHV, the original intent was to create a fully structured multi-block volume grid using Link3D. Using this approach to make a quality 3D structured grid was proven to be unattainable due to the complexity of the GHV geometry. Specifically, the surface topology tended to cluster a large number of cells inside the vehicle relative to the outside causing a sharp disparity of cell spacing. This caused the grid smoother in Link3D to pinch the cells in the inlet region which rendered the grid unusable. After many failed attempts to balance the cell spacing, a different approach was taken to generate the 3D grid.

In an effort to create a usable, quality grid, a hybrid 3D grid was generated using Pointwise. Pointwise is a commercial gridding software that excels in creating unstructured and hybrid meshes. The structured surface grid from Link3D was imported into Pointwise and used to define the GHV wall. The farfield boundary was chosen to be a half-cone with a hemispherical tip that closely fits around the Mach cone emanating from the vehicle. Note, the grid cone angle was set to 25.0° while the Mach cone angle from the lowest Mach number (3.0) had an angle of 19.1°. The boundary used a Delaunay unstructured scheme that used triangles to define it. A tight farfield boundary around the surface is allowable when simulating hypersonic flow because information can only propagate downstream from shock waves meaning upstream cells beyond the Mach cone are not influenced by the vehicle surface aerodynamics [12]. This occurs because the speed of the airflow prevents any interactions or effects from the vehicle to propagate outside the Mach cone.

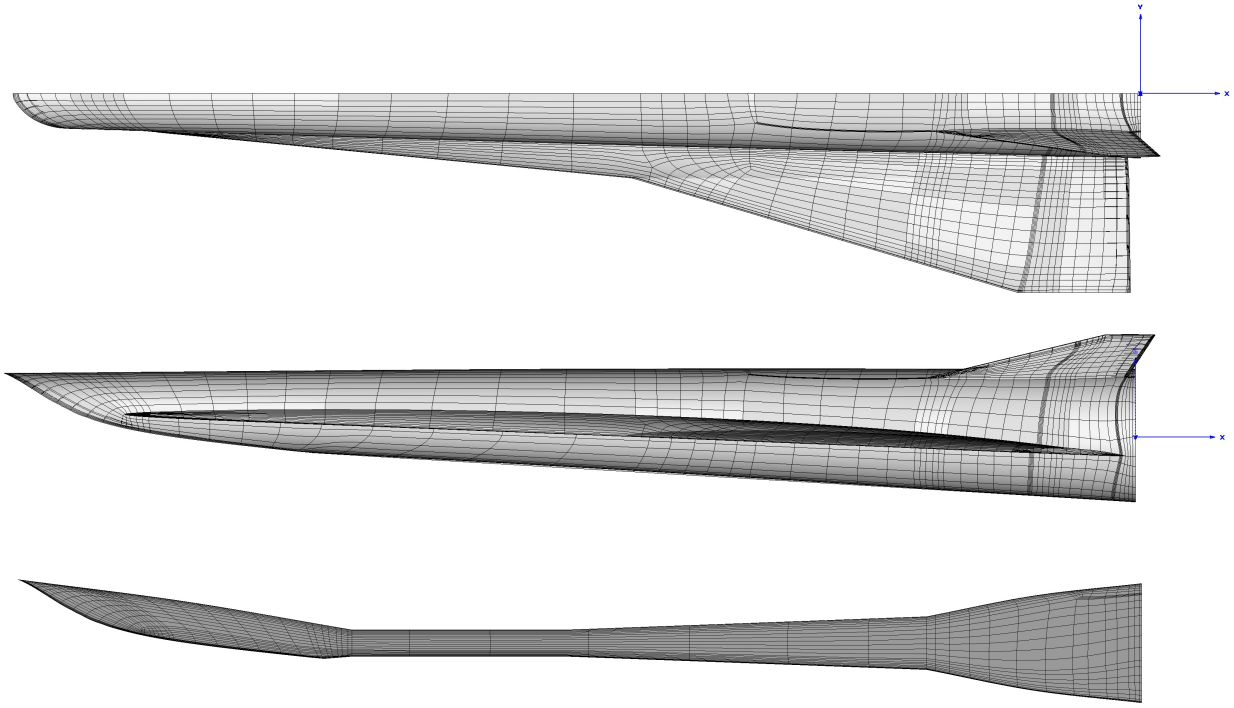


Figure 16: Surface Grid (Coarsened for Visibility)

Layers of structured Anisotropic (T-Rex) cells were created surrounding the vehicle surface. The purpose of using T-Rex cells was to evenly model the flow surrounding the surface. For the inviscid Euler grid, the T-Rex cells did not need to be highly resolved around the surface because the boundary layer is not modeled. Thus, the initial T-Rex cell spacing was set to 1 mm (1×10^{-3} m) with a growth rate of 1.2 with 10 cells in the layer. After the creation of the T-Rex layer, isotropic cells were computed for the rest of the flow domain using the Delaunay algorithm and a decay rate of 0.75. The symmetry plane was simply specified to match the T-Rex cell dimensions and isotropic cells in the volume (fig. 17). Figure 18 shows the overall computational domain including the wall boundary (gray), the farfield boundary (blue), and the symmetry plane (red). Figure 19 visualizes the volumetric cells surrounding the vehicle via cut planes. In the figure, tetrahedrons are red, pyramids yellow, and rectangular prisms (T-Rex) blue.

Throughout this research, a common coordinate system was used. The origin was defined at the point (0,0,0) and was located at the rear of the vehicle, coincident with the center of the nozzle opening. The x axis pointed aft of the origin parallel to the length of the vehicle. Thus, the front-most point of the GHV was located at $x = -4.318$ m. The y axis pointed from the origin to the right wing. Only the -y half of the vehicle was modeled because of the symmetry plane. Finally, the z axis pointed from the origin to the top of the vehicle. All point locations reported

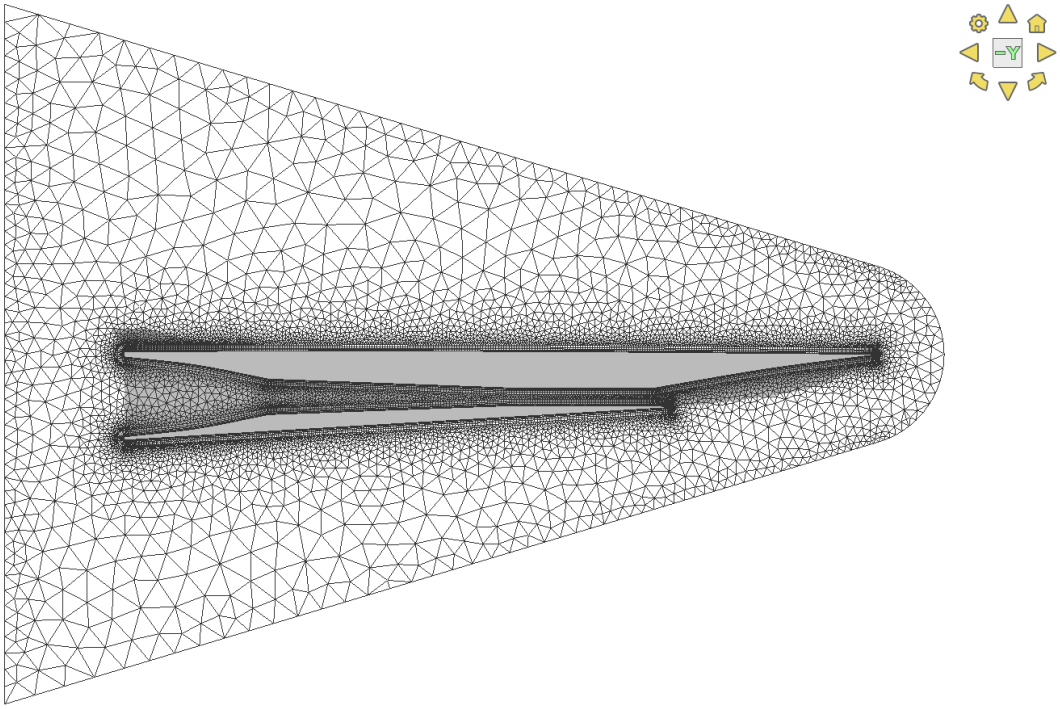


Figure 17: Symmetry Plane Featuring T-Rex Cell Layers

hereafter will follow this coordinate system.

4.4 Grid Convergence Study

In order to build an accurate surrogate model of the aerodynamic responses of the GHV, the training data itself needed to have minimum error. One source of CFD error stems from the nature of the computational grid. The purpose of a grid convergence study is to prove that the CFD data is *grid independent*, meaning that the data remains unchanged after continued grid refinement [12]. This is accomplished by increasing the number of grid cells until the flow properties of interest are no longer affected.

It is important to conduct the convergence study using the parameters of interest. For instance, integrated forces such as C_L and C_D tend to converge more rapidly than a more sensitive parameter like a shock wave location or a pressure value at a specific point. If the research is only interested in the integrated forces, then a coarser grid may be selected. However, it is important to note that while this grid may be accurate for integrated forces, it may not provide accurate results for other, more sensitive parameters [12]. In this research, the primary parameters of interest are integrated forces and surface pressure values, thus the convergence study will assess the convergence of these values.

A set of 10 grids with varying cell counts were tested under four common freestream conditions and compared to each other. The grids consisted of resolutions ranging from 70 thousand up to 25

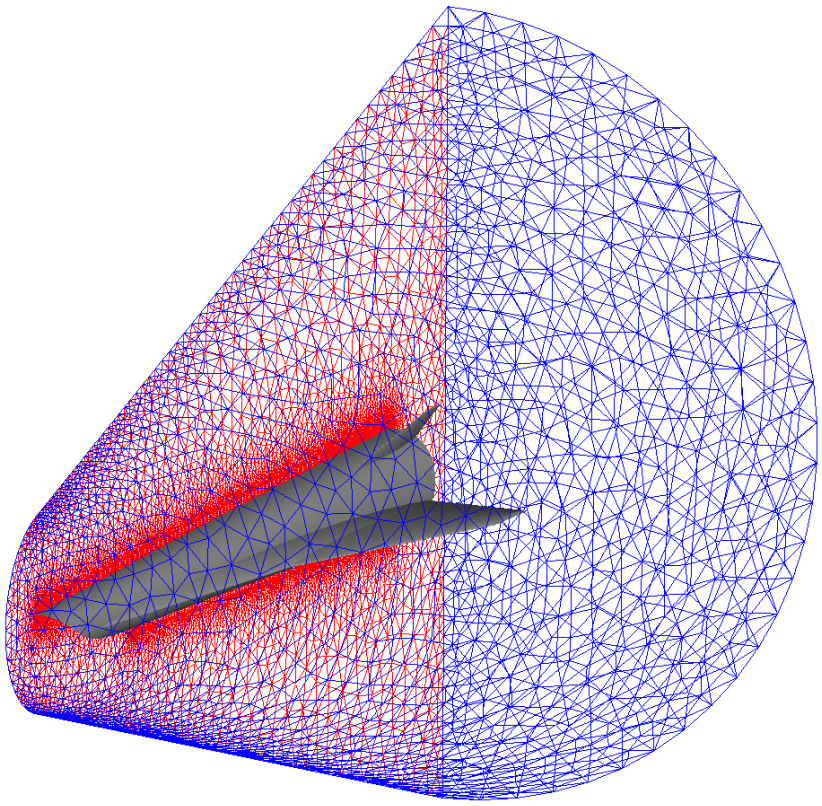


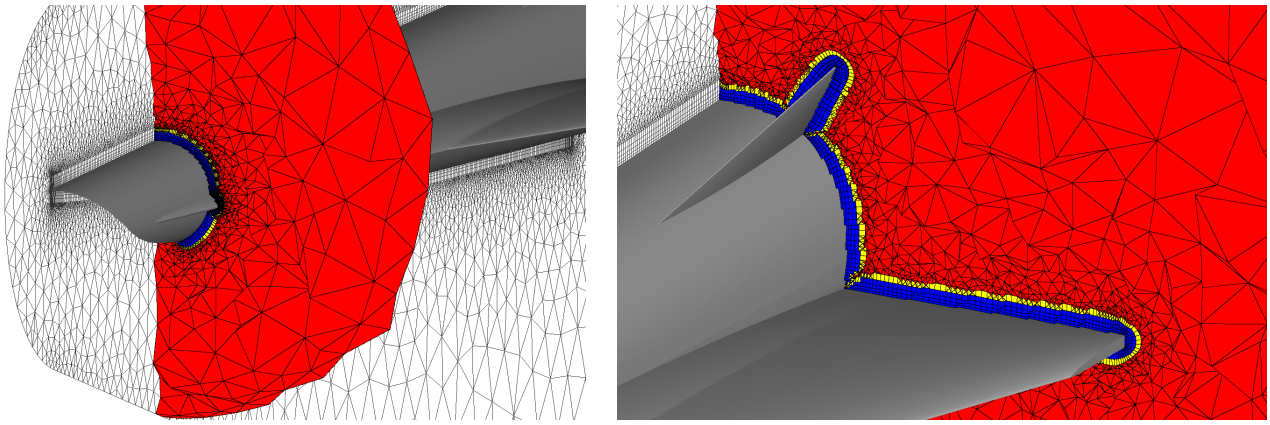
Figure 18: Volume Grid Boundaries

million cells as shown in Table 5.

Each grid was refined globally, but with a focused refinement on the surface grid. This was accomplished by first specifying the refinement of the structured, surface grid by adjusting the grid spacings by a factor of $\sim$1.2 in Link3D. This in turn increased or decreased the surface cell count by approximately $(1.2)^2 \rightarrow 1.44$ times because of the surface grid being two-dimensional. The unstructured volume grid connector spacing was edited manually in Pointwise to increase the global refinement throughout the computational domain. After the refinement process was completed, CFD simulations were executed on each grid for comparison.

The purpose of using CFD in this research was to collect computational data across a design space of varied flight conditions, thus the CFD grid needed to be verified and converged throughout this range of conditions in order to provide accurate data. Thus, four different Test Conditions (TC) were specified in order to represent the design space defined in Table 1. Simulations were run at these test conditions for each of the 10 grids. Table 6 lists these test conditions.

The parameters assessed for the convergence study were integrated quantities (C_L, C_D) and surface pressures. These particular parameters were assessed to provide justification for the



a) b)

Figure 19: Volume Grid Cut Plane a) Front Fuselage b) Rear Fuselage

validity of the training data used later for the surrogate model construction. Using the C_L and C_D computed from the Kestrel Euler CFD code, the grid convergence was assessed at each test condition as shown in Figure 20. Note, the C_L and C_D for the last 100 iterations were averaged for each run. The four subplots in Figure 20 plot C_L and C_D computed from the corresponding grid's cell count. Note the cell count is plotted on a logarithmic axis.

For each test condition, the C_L and C_D values tend to converge to a finite value. This shows that the values of C_L and C_D become less affected by the cell count as the cell count increases. Note that each of the trends do not converge monotonically, but rather in a random, erratic manner with decreasing amplitude. This is most likely due to the grid being unstructured. During the grid generation process, the Delaunay algorithm is used for the placement of tetrahedrons in the volumetric domain and is dependent on the boundary surface grids (farfield, symmetry, and wall). Because the grid is refined globally via refining the GHV surface grid and the farfield/symmetry plane connectors, the exact location of the tetrahedrons determined by the Delaunay algorithm differs randomly from grid to grid, causing seemingly random variations in the grid. These variations are most likely the cause of the "random" convergence behavior observed in the integrated forces.

Next, the surface pressures were analyzed for convergence. Four points spread across the GHV surface were selected as reference locations: the upper wing, the lower wing, vertical tail side, and the upper-front part of the fuselage. These surface point locations are shown in Figure 21. The pressures were extracted using a simple 3D interpolation using data from the final CFD iteration. Pressures were not extracted from individual points because each grid had different surface point placements due to the differing refinements, thus an interpolation was needed to ensure a common surface location. The pressure convergence plots at these four points for TC-1 are plotted in Figure 22. These trends show similar random variations exhibited in the trends in Figure 20. The upper wing and front-upper fuselage showed the most striking convergence trends.

Table 5: Euler Grid Convergence Cell Counts

Grid ID#	Description	Total Cell Count	Surface Cell Count
1	Extremely Coarse	73,292	3,044
2	Very Coarse	350,453	13,680
3	Coarser	850,641	34,827
4	Coarse	2,609,842	98,426
5	Medium	3,350,606	127,421
6	Fine	4,960,274	176,825
7	Finer	8,777,642	247,989
8	Very Fine	9,939,874	347,863
9	Extremely Fine	15,698,233	490,922
10	Superfine	24,910,775	803,465

Table 6: Grid Convergence Test Conditions

Test Condition	M	α (°)	h (m)
TC-1	6.0	0.0	25,000
TC-2	3.5	2.0	15,000
TC-3	5.0	5.0	20,000
TC-4	7.5	-3.0	27,000

The other two trends were less satisfactory, but they generally converged to a specific pressure. Plots were generated for the other three TCs and are plotted in Figures A.2 to A.4 within Appendix A. They show similar trends as those discussed in Figure 22.

Finally, using the information gleaned from the grid convergence study, a grid was selected. The final grid was chosen qualitatively due to the complexity of the convergence trends. If all of the trends were to be summarized, the parameter's variation starts very large in amplitude for the coarser grids, then decreases in magnitude as the grid is refined. Observing the data, the range of magnitude from grids 6 to 10 was smaller. The magnitude of these variations were estimated using a percent difference formula shown in Equation (41) where $\mathbf{y}$ is the set of parameters (C_L, C_D, or P_{surf}) from grids 6 to 10.

$$\%D = \left| \frac{\max(\mathbf{y}) - \min(\mathbf{y})}{\bar{\mathbf{y}}} \right| * 100\% \tag{41}$$

The percent differences of each parameter and test condition is shown in Table 7. From this we

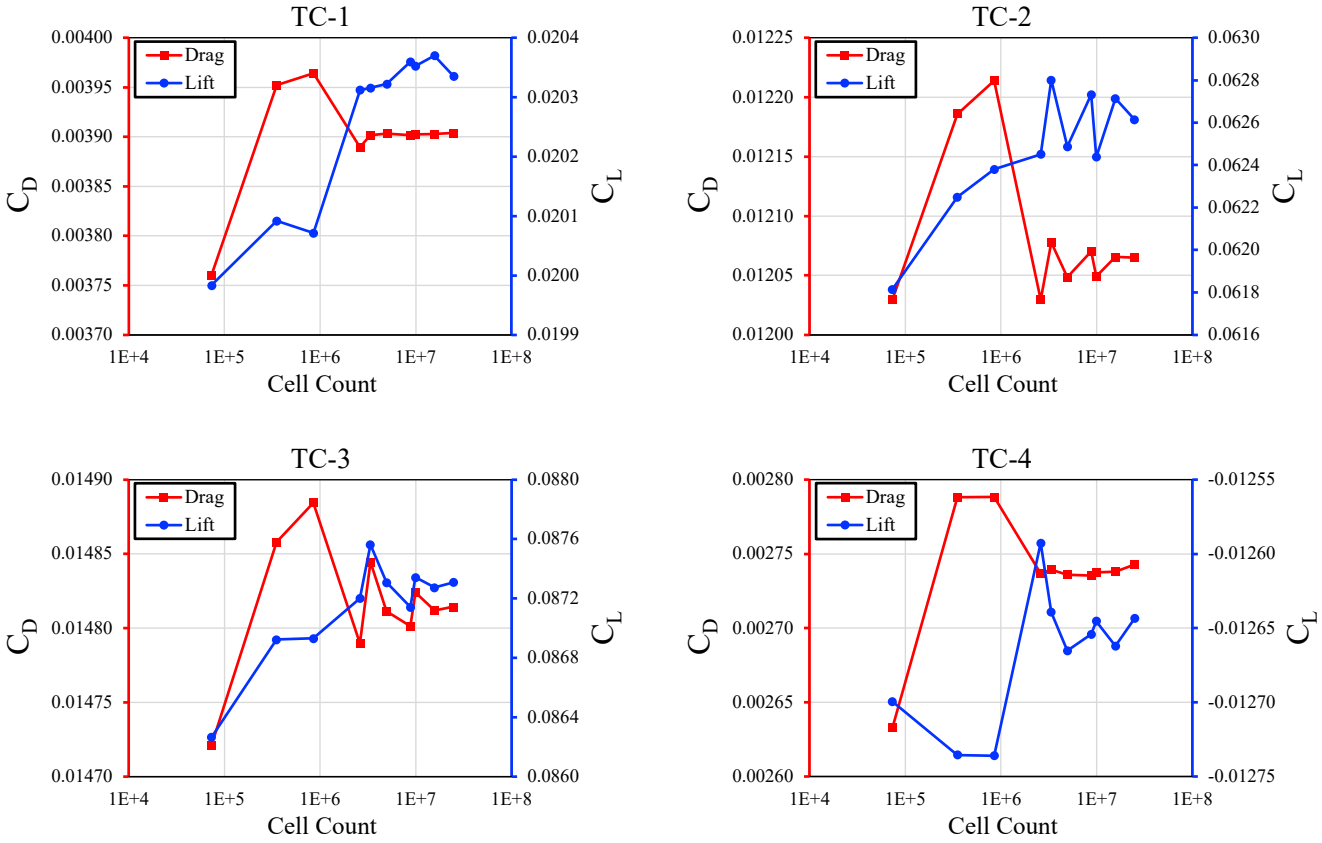


Figure 20: Euler Grid Convergence of C_L and C_D for Each Test Condition

see that the integrated force percent differences were all well under one percent. As expected, the surface pressures performed more poorly with a percent difference of around 5% or less (with the exception of P_2 at TC-4).

Inspecting the trends in Figures 20 and 22 and understanding the potential errors shown in Table 7, Grid# 7 with 8.8M cells (finer grid) was selected as the final grid for the training data collection. It was determined that the parameters of interest were appreciably converged at this grid resolution.

4.5 Data Collection

4.5.1 Sampling Strategy

To provide the optimal placement of points throughout the design space, the quasi-random Sobol set was selected. This choice was made because the Sobol set maintained low discrepancy (good space-filling) of points regardless of the size of the set, and this feature was needed for building different sized surrogate models later on. Refer to Section 3.5 for more details on the Sobol set.

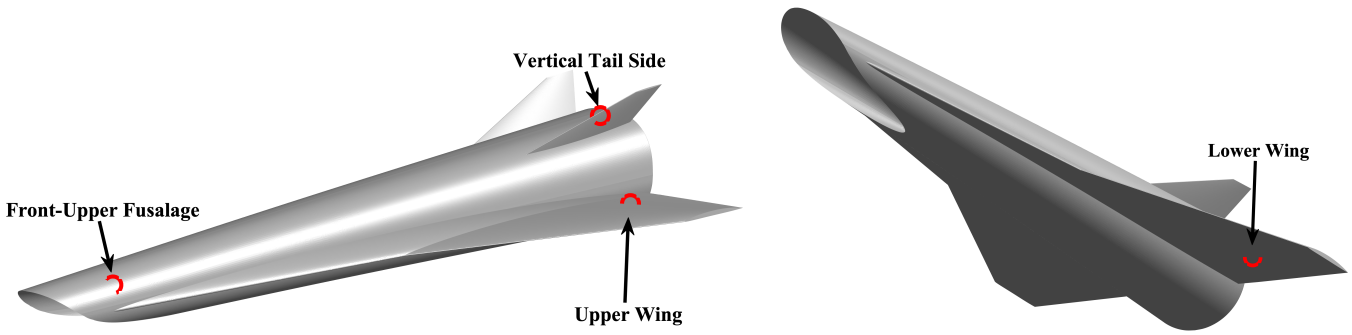


Figure 21: Convergence Surface Pressure Locations

Table 7: Maximum Percent Differences for Grids 6 to 10

Test Condition	C_L %D	C_D %D	P_1 %D	P_2 %D	P_3 %D	P_4 %D
TC-1	0.06	0.23	1.53	4.20	2.05	0.16
TC-2	0.18	0.37	5.04	1.93	1.95	0.38
TC-3	0.15	0.18	4.57	3.20	3.63	0.32
TC-4	0.26	0.12	1.76	14.89	1.11	0.69

The three-dimensional design space boundaries are defined again in Table 8. A Sobol set of 1000 data points was generated using these boundaries and used as the CFD inputs for the training data. These 1000 data points are plotted in Figure 23.

Table 8: Sample Design Space

Parameter	Min	Baseline	Max
M	3.0	6.0	8.0
α (°)	-4.0	0.0	6.0
h (m)	15,000	25,000	27,500

4.5.2 Flow Solver

The CFD solver used for this study was the CREATE-AV Kestrel software package. Kestrel version 11.1.2rc2 was consistently used for all of the CFD data discussed in this document. The solver used an implicit, steady-state scheme to compute the solution for each run. The Euler equation was used for the training data. Although Euler does not model viscous effects of the air, it runs an order of magnitude faster than RANS. The HPC computer resources and collection time required to use RANS for the surrogate training data was not available during this research, which is why Euler data was collected.

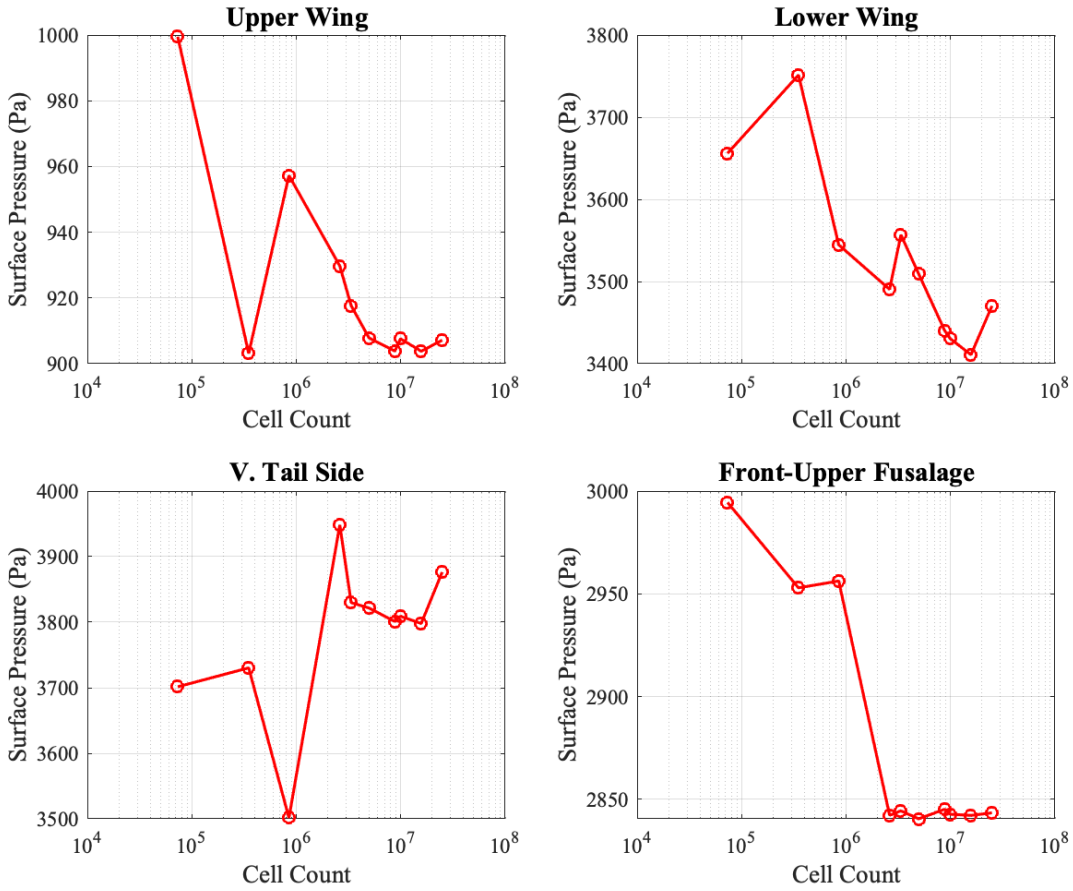


Figure 22: Euler Grid Convergence of Surface Pressures for TC-1

Each data point ran to 3000 iterations which was determined by inspection of residual plots. For the typical data point, the CFD solution converged before 1000 iterations, so it was a safe assumption that the simulations were well converged by the 3000 iteration mark. An example residual plot for RUN-0001 is shown in the appendix in Figure A.5. The CFL number was initialized at 0.1 for the first iteration, then it was ramped up to 100 for the 1000th iteration, and it remained constant for the remainder of the simulation. It was found that the solution was negligibly dependent on the CFL number. This was determined from a brief CFL number analysis which will be discussed in Chapter 5.

In order for Kestrel to calculate the force and moment coefficients, a set of reference parameters needed to be specified. The reference area was set to 1.6375 m^2 which is half of the GHV's planform area (note: the grid only models half of the vehicle using a symmetry plane). The center of mass was assumed to be located at the geometry's centroid at (-1.68,0,0.0674) meters. This assumption only affects the values of the moment coefficients. Finally, the reference length for the moment calculations was set to the length of the vehicle at 4.318 m.

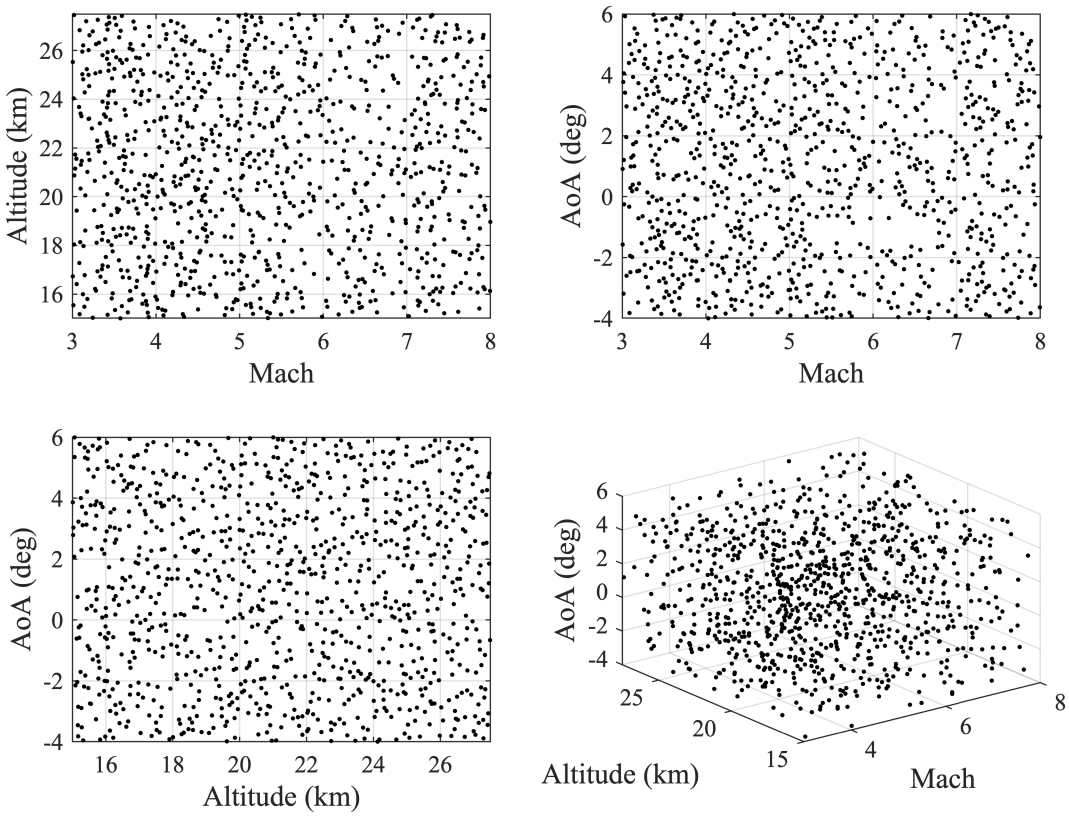


Figure 23: Sample Point Distribution

The freestream conditions (Mach, α, and altitude) for each run was determined by the sequence of points generated using the Sobol set. The Mach number and angle of attack values were entered as-is for the solver settings. Kestrel used the 1976 U.S. Standard Atmosphere Model [34] to convert the altitude (in meters) into the appropriate ambient pressure and temperature. All other solver setting were held constant.

4.6 Data Reduction and Preparation

After all of the CFD simulations finished running, certain parameters of interest were extracted from the solution. In general, they consisted of integrated quantities and surface point quantities at certain critical points. Table 9 shows a list of the parameters of interest. Seven critical points representing areas of large surface pressure and heat flux were selected for the surrogate analysis based on inspection of the surface pressure distribution. The selection of these points is explained in detail in Section 5.2.

Each of these parameters were extracted from the Kestrel output files. C_L, C_D, and C_m were

Table 9: Surrogate Parameters of Interest

	Symbol	Description	Units
Integrated	C_L	coefficient of lift	-
	C_D	coefficient of drag	-
	L/D	lift to drag ratio	-
	C_m	pitch moment coefficient	-
Surface	P_{crit}	surface pressure at critical point	kPa
	P_{max}	max surface pressure	kPa
	P_{min}	min surface pressure	Pa
	h_{crit}	heat flux coefficient at critical point	$\frac{\mathrm{W}}{\mathrm{K\ m^2}}$
	$\dot{q}_{crit}$	heat flux at critical point	$\frac{\mathrm{kW}}{\mathrm{m^2}}$

extracted by taking the parameter's respective means from the past 100 iterations of the CFD solution. This was done in order to remove the numerical noise from the converged values of C_L, C_D, and C_m. These values were found in the coeff.dat file in the Kestrel output. The lift to drag ratio, L/D, was calculated by simply dividing C_L by C_D.

In order to find the surface values, a Tecplot macro was used to take the binary visualization file from Kestrel and extract all the surface points and their parameters as a comma-separated-value file. The maximum and minimum pressures were found from this file and their respective locations were captured. Then the critical points were located and the pressure, temperature, Mach, and airspeed at these points were extracted. These were used to calculate the heat flux coefficient (h) and the heat flux ($\dot{q}$) values via ERT. The heat flux calculation used a wall temperature of 300 Kelvin consistently throughout the analysis.

After the process of extracting and calculating parameters was complete, all 1000 samples of varied Mach, AoA, and altitude had corresponding output data. This data was used as the input for the kriging software and was used as training data to build the surrogate models.

4.7 Kriging Surrogate Construction

The software used to generate surrogate models from the CFD data was ooDACE [22], which stands for Object Oriented Design and Analysis of Computer Experiments. ooDACE is a MATLAB based software that uses kriging and other methods to create surrogate models. It utilizes a third-party optimization software called SQPLab in order to solve the maximum likelihood estimation. ooDACE contains several different options for correlation functions: cubic, exponential, Gaussian, linear, Matern-32, Matern-52, spherical, and spline. The software package also allows for the specification of the regression polynomial order ranging from 0 to 4th order. These settings were tested and evaluated in Section 5.3 so the optimal correlation function and regression order can be found.

5 RESULTS AND ANALYSIS

5.1 CFL Number Study

Before the training data was collected for the surrogate model, a short study was conducted to determine the effect of the CFL number on the CFD solution. The CFL number, or *Courant* number, named after the discoverers, Courant, Friedrichs, and Lewy [35], describes the relation between the spacial step size and the time step size and dictates the stability of an explicit integration scheme [12]. The expression for the CFL number is given in Equation (42) and c is equal to the wave speed, Δt is the time step, and Δx is the spacial step.. Since we are using implicit integration, the solution should be unconditionally stable with respect to the CFL number. For a steady state simulation, the time order is not significant, and the solution propagates the time until a steady solution is obtained [12].

$$CFL = c\frac{\Delta t}{\Delta x} \tag{42}$$

The question that needed to be answered was whether the specified CFL number had an impact on the outcome of the solution, and if so, which CFL number provided the most stable results. In theory, a lower CFL number results in a more stable solution and quicker computation time per iteration due to the lower time step, but it also increases the number of iterations required to converge to steady state.

Eight different tests with varying CFL numbers were run at common freestream conditions. The Mach number was set to 6.0, angle of attack to 0.0°, and altitude to 25,000 m. Each run was iterated until convergence, which was confirmed by reviewing the residual histories. The C_L and C_D of each run is provided in Table 10 along with the convergence iteration and the time required to reach convergence in processor-hours.

Table 10: CFL Number Test Results

CFL	C_L	C_D	Iterations	Time to convergance (CPU-hrs)
0.5	0.008362925	0.001606441	13500	391.1
1	0.008365091	0.001606647	7000	208.1
5	0.008368818	0.001606510	1900	56.93
10	0.008368151	0.001606628	1200	24.77
50	0.008365802	0.001606622	800	24.50
100	0.008364104	0.001606417	700	25.25
1000	0.008363691	0.001606455	600	23.89
2000	0.008363332	0.001606439	600	23.85

The C_L and C_D varied at a maximum of 5.9×10^{-6} and 2.3×10^{-7}, respectively. This indicated that there were no solution dependencies on CFL number. However, the number of iterations

needed for convergence decreases with an increase in CFL. Additionally, for very low CFL numbers 5 and under, the simulation time from start to convergence is very large. CFL numbers 100 and above do not show any major improvement to the number of iterations and time required to reach convergence. With these results in mind, a conservative CFL number of 100 was chosen for the flow solver settings throughout the research.

5.2 Qualitative CFD Analysis

Before building surrogate models, the pressure and Mach profiles of the flowfield were qualitatively analyzed. Additionally, the surface pressure contours on the GHV were inspected and critical points of large pressures were selected. The pressures at the selected critical points were modeled later using Kriging and the surface heat fluxes were calculated later using ERT.

5.2.1 Flowfield Analysis

One of the most comprehensive views of the flowfield pressure distribution is the symmetry plane. The converged GHV grid was computed using the settings discussed in Chapter 4 at the nominal flight condition: $M = 6$, $\alpha = 0°$, and $h = 25,000$ m. Figure 24 shows the pressure contours of the symmetry plane on a logarithmic scale at these conditions. Note, the symmetry plane cuts through the entirety of the fuselage which is displayed in white. As expected, the pressure was constant outside of the Mach cone due the hypersonic flow preventing information propagating upstream. The Mach cone shock waves were seen emanating from the leading edge of the inlet and expanding above and below the vehicle. The lower shock was noticeably stronger than the top because of the isentropic inlet wall's large incidence angle relative to the flow. Near the lower inlet lip, there was some *spillage* [10] of the flow indicted by the release of high pressure flow below the vehicle outside the scramjet.

The scramjet flowpath was clearly visible in this view, and some interesting features were observed. The isentropic inlet progressively compressed the airflow prior to entering the scramjet, which validated its geometric design. A key feature noticeable in Figure 24 is the oblique shock-train starting from the lower inlet lip and ending inside the combustor. This confirms the functionality of the isolator. Throughout the isolator and the beginning of the combustor, the air pressure was especially high. At this particular flight condition, the compression ratio, π_c, was approximately 22.0 at the beginning of the combustor. The scramjet nozzle significantly lowers the pressure due to the expansion of the flow. The low pressure aft of the nozzle induced acceleration of the flow causing high Mach numbers. Overall, the Euler CFD showed that the unstarted scramjet operated as expected.

Next, the symmetry plane was plotted alongside 3 different Y-Z cut-planes in Figure 25 to provide a three dimensional view of the flowfield pressure. This view shows the Mach cone growing in diameter as the airflow progresses downstream. Because of the inclined inlet face and the inclination of the wing, the pressure was larger below the GHV which indicated that the vehicle was producing lift at these conditions. The GHV's strake and wing assisted in maintaining the

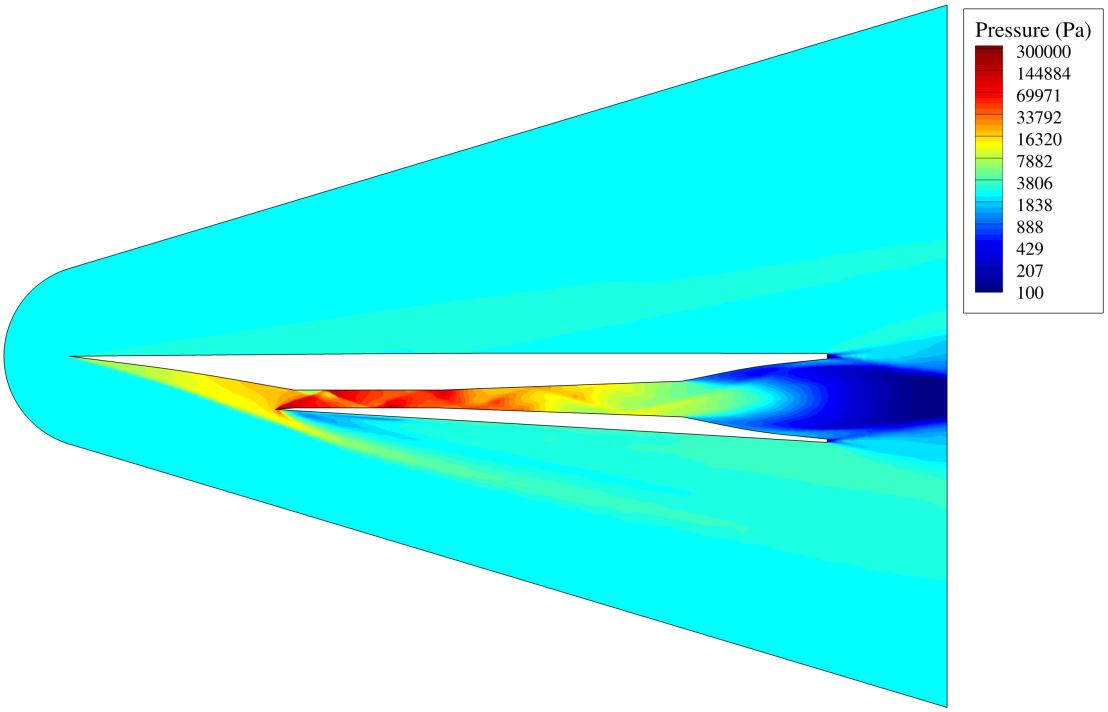


Figure 24: Symmetry Plane Pressure Contours, $M = 6.0$, $\alpha = 0.0°$, $h = 25$ km

high pressure below the vehicle. The GHV could be considered a wave-rider because it uses the high pressure generated from the inlet's shock waves to produce lift. Inspecting the wing near the trailing edge, the pressure difference between the top and bottom wing caused the flow to circulate causing vortices. This is evident due to the spill-over of high pressure air from the bottom of the wing to the top at the tip of the wing.

Figure 26 shows the pressure contours on a X-Z cut-plane at a constant Y=0.6 m value which displays the cross section of the wing. The wing generates a prominent bow shock wave at the leading edge which results in very high pressures. The pressure dropped significantly above the wing due to the expansion of the flow from the cambered biconvex airfoil. The wing incidence angle caused higher pressures below the wing, which then generated lift from the top-to-bottom pressure difference.

Finally, the Mach number contours were shown in Figure 27 on the symmetry plane at the nominal flight conditions. Starting upstream, the first change in Mach number began at the inlet due to the strong oblique shock. The flow gradually slowed down throughout the inlet due to the isentropic compression. The airflow slowed down further after each consecutive oblique shock in the shock-train inside the isolator. As expected, the flow was accelerated in the combustor and nozzle due to the expansion of the flow and the reduction of pressure. Aft of the nozzle, the flow

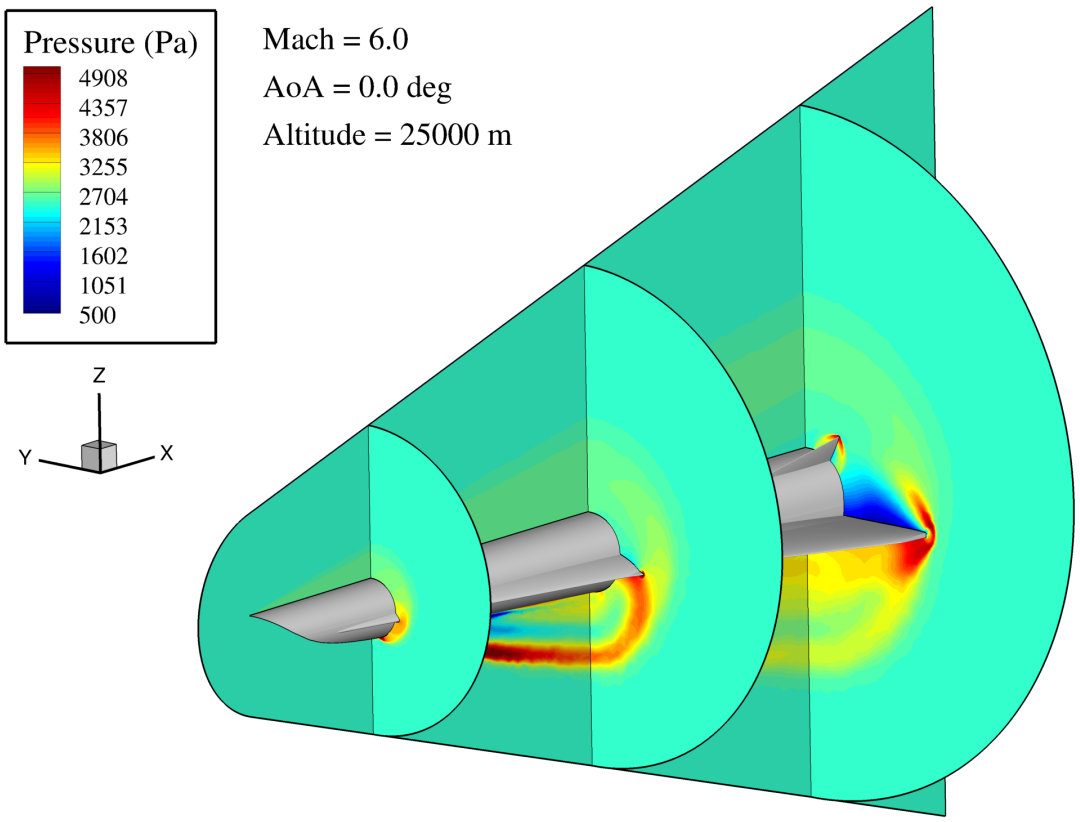


Figure 25: Pressure Contours on Symmetry Plane and 3 Y-Z Cut-Planes, $M = 6.0$, $\alpha = 0.0°$, $h = 25$ km

accelerated up to Mach 10.

5.2.2 GHV Surface Analysis

The pressure contours on the surface of the GHV were plotted using four views: top, bottom, side, and the interior of the vehicle which are shown in Figure 21. Note, the computed surface pressures were mirrored at the symmetry plane to display the entire vehicle. The highest pressures across the entire surface were located inside the vehicle at the inlet throat and isolator region. Outside of the vehicle, the highest pressures were located on the leading edges of the inlet, wing, and vertical tail.

Seven critical points were selected from these regions of high pressure throughout the vehicle. The flow conditions at these points were extracted later from each run and used as training data for the surrogate model construction. The critical points and their locations are specified in Table 11. The critical point locations are specified by the magenta circles in Figure 28. It is assumed that these points also represent areas of large heat flux because ERT states that the heat flux is

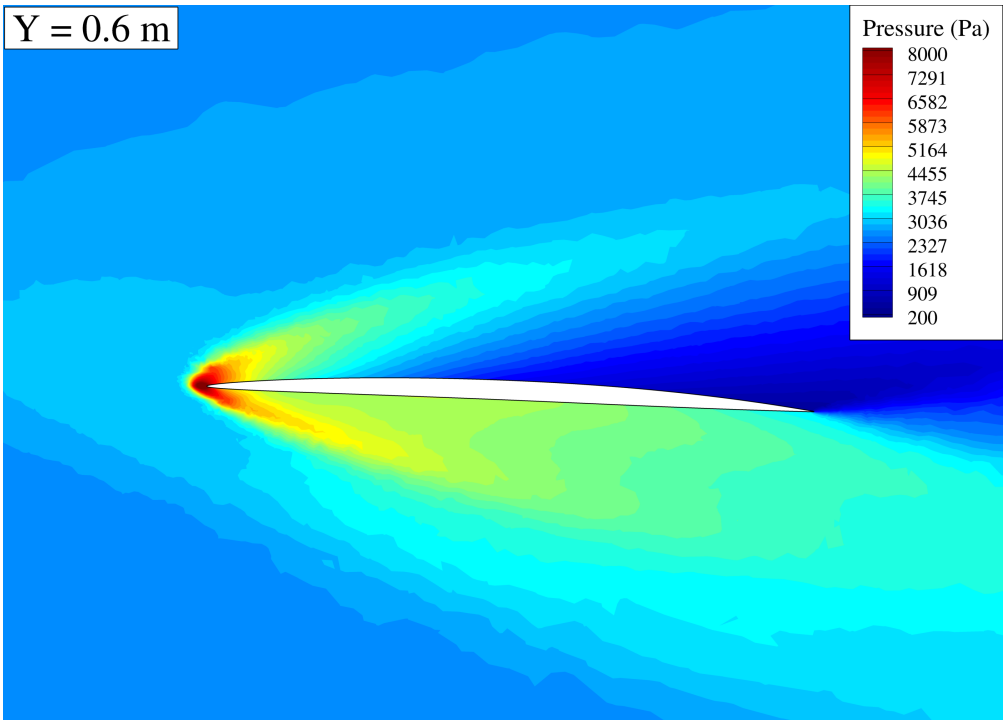


Figure 26: Pressure Contours on X-Z Cut-Plane, $M = 6.0$, $\alpha = 0.0°$, $h = 25$ km

proportional to the local surface pressure. Refer to Section 3.3 for a more thorough explanation.

5.3 Surrogate Correlation and Regression Function Analysis

An analysis comparing the different correlation functions and regression functions was conducted to determine the optimal settings for the surrogate construction. A total of eight correlation functions were tested at each of the five regression orders ranging from 0th to 4th order. Surrogates were constructed for twelve output parameters using these correlation and regression functions. Every combination of output parameter, correlation function, and regression function was tested which amounted to 480 surrogates. Table 12 lists all of the parameters and functions used for the surrogate tests.

A reasonably large set of 300 points was specified as training data and a different set of 50 points were used as validation data with the purpose of evaluating the error of the surrogates. The point distribution is shown in Figure 29. All 480 different cases were run using the same 300 training points and predictions were made at the locations corresponding to the validation data. The prediction and validation points were used to calculate the Root Mean Square Error (RMSE). The equation for RMSE is,

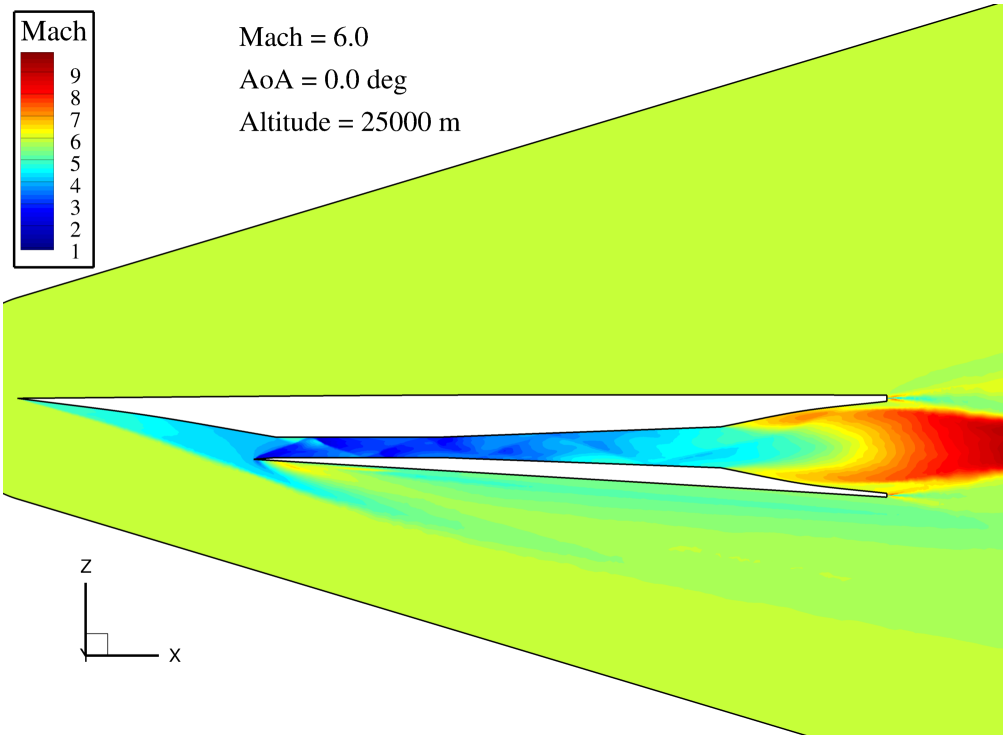


Figure 27: Symmetry Plane Mach Contours, $M = 6.0$, $\alpha = 0.0°$, $h = 25$ km

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \qquad [36] \tag{43}$$

where y_i is the ith validation point, $\hat{y}_i$ is the ith prediction point, and N is the total number of validation points. Note that the RMSE is dimensional; it contains units equal to the original parameter. Thus, two RMSE values from different variables cannot be compared directly due to the difference in units and/or the difference in magnitude.

The lowest (optimal) RMSE from the 5 regression orders was selected for each parameter and correlation function and displayed on Table 13. The best correlation function RMSE was highlighted in green for each parameter. Entries with *NaN* signify that a critical failure occurred during surrogate construction for each of the 5 regression functions. Failures occur due to the kriging correlation matrix failing to be positive definite, and this is dependent on which correlation function is used and the particular distribution of training points.

Inspecting Table 13, we see that the cubic function failed regularly which is an indication that it should not be used further on. The exponential and Matern-32 correlation functions performed better overall. For most of the critical pressures, some correlation functions resulted in errors an

Table 11: Surface Critical Point Locations

Point #	x	y	z	Description
1	-4.31508	-0.005065	0.242802	upper inlet lip
2	-3.14332	-0.005515	-0.058905	lower inlet lip
3	-3.85222	-0.140321	0.087342	fuselage/strake junction
4	-2.54867	-0.265915	0.034053	strake leading edge
5	-1.24592	-0.532030	-0.021296	wing leading edge
6	-0.48197	-0.188675	0.302118	vertical tail leading edge
7	-3.03530	-0.008879	0.050027	lower inlet throat

order of magnitude worse than the optimal RMSE. For example, the best correlation function for P_{max} was exponential with an RMSE of 178, while the worst was the linear correlation function with an RMSE of 7756, which is larger by a factor of 44. This shows the importance of selecting the correct correlation function when building surrogate models to maximize the accuracy of the predictions.

A summary of the optimal correlation function and regression order is shown in Table 14. These were selected based on the RMSE calculated from each of the settings tested. It is interesting to note that C_L and C_D preferred a 1st order regression polynomial meaning the overall trend was roughly linear. On the other hand, the critical pressures typically preferred 3rd and 4th order regression polynomials to model the supposedly more drastic changes across the design space. While this does provide useful information on surrogate settings for individual variables, it does not provide much insight for optimal surrogate settings for overall performance across multiple parameters.

To determine the best settings for the overall surrogate performance, the RMSEs were ranked from low to high and a composite score was calculated. This was done for both the correlation functions and the regression functions. The correlation functions were ranked from 1 to 8 based on their RMSE values which is shown in Table 15. After the functions were ranked, a composite score was taken using three bins: integrated quantities containing C_L, C_D, and C_m; surface quantities containing all critical pressures, and total which contained all 12 parameters. The score was calculated by taking the mean of the ranks in each respective bin and each correlation function. The lower the score, the better the correlation function performed for that particular bin.

It was found that the Matern-32 correlation function performed best with the integrated quantities (C_L, C_D, and C_m) and overall for all parameters. The Gaussian correlation function performed best while modeling surface quantities. Due to the overall performance of the Matern-32 correlation function, it will be used for the future analyses moving on.

A similar analysis was completed for the regression functions. The 0th to 4th order regression

Table 12: List of Parameters, Correlations, and Regressions Tested

Parameter	Correlation Function	Regression Order
C_L	Cubic	0th
C_D	Exponential	1st
C_m	Gauss	2nd
L/D	Linear	3rd
P_{crit1}	Matern-32	4th
P_{crit2}	Matern-52	
P_{crit3}	Spherical	
P_{crit4}	Spline	
P_{crit5}		
P_{crit6}		
P_{crit7}		
P_{max}		

function RMSEs were ranked from 1 to 5 for each parameter in Table 16 and composite scores were calculated for the same bins described above. Since the composite scores in the ”total” bin are similar to each other, this indicates that there is very little agreement on which regression order is optimal overall. Thus, when selecting a surrogate regression function for a variety of variables, the importance of which regression function is selected is lower. The best overall regression was the second order polynomial which didn't score first with any of the parameters, but rather performed proficiently for each. However, for integrated quantities, a linear regression order (1st order) is clearly optimal and should be selected when modeling these values.

Finally, a table showing the cases where there was a failure in the surrogate construction was made which is shown in Table 17. Each cell in the table sums the number of failures from the 5 regression function cases for each parameter and correlation function. It is clear that the cubic correlation function was the most unstable because of the high likelihood of failure. Additionally, some failures occurred while using the Gaussian and Matern-52 functions. If the stability of the surrogate model construction is important, the cubic, Gaussian, and Matern-52 correlation functions should be avoided.

Several conclusions can be made from the results of this study of surrogate generation settings. If the highest possible accuracy is needed when predicting a particular parameter, a tailor-made regression function and correlation function needs to be selected based on the model's error. Each parameter typically preferred a unique combination of correlation function and regression order. If many parameters need to be modeled with a common surrogate, the Matern-32 correlation function with a 2nd order regression polynomial is a good option. The Matern-32 correlation is

Table 13: RMSE Values for Different Surrogate Correlations and Parameters using 300 Training Points and 50 Validation Points

Parameter	Cubic	Exponential	Gauss	Linear	Matern32	Matern52	Spherical	Spline
C_L	0.000018	0.000005	0.000023	0.000042	0.000003	0.000015	0.000003	0.000007
C_D	0.000164	0.000007	0.000164	0.000095	0.000164	0.000081	0.000087	0.000164
C_m	0.1690	0.0555	0.0323	0.0124	0.0019	0.0021	0.0699	0.0045
L/D	0.1206	0.0295	0.0162	0.0180	0.0077	0.0041	0.0105	0.0183
P_{crit1}	43.0	170.9	5.43	134.7	58.6	25.6	68.0	48.2
P_{crit2}	NaN	3861.8	99.4	11.6	655.5	3754.1	2560.2	3492.4
P_{crit3}	NaN	19.5	1.47	9.58	0.84	0.78	14.3	1.20
P_{crit4}	NaN	201.8	13.6	193.2	157.6	136.4	191.1	89.8
P_{crit5}	10.6	6.07	3.79	37.2	1.91	9.12	46.2	23.1
P_{crit6}	98.4	12.8	37.5	213.6	77.2	83.5	62.8	3.84
P_{crit7}	1774.9	36.5	1774.9	647.6	67.8	175.6	1053.7	863.1
P_{max}	NaN	178.5	346.8	7756.1	5913.9	3001.7	2506.9	2616.0

also desirable because there were no critical failures in any of the 60 cases tested with it. It is important to stress that there are other capable surrogate model options that fit the data well. Discernment on the user's part is needed in order to select the best settings for the particular application.

5.4 Surrogate Model Sizing

One of the most important questions about the development of a Kriging model is assessing how many training points are needed to create a sufficiently accurate model. More training points result in a more accurate surrogate, but if the training data is expensive to obtain, a balance needs to be struck between maximizing the accuracy of the model while minimizing the number of training points. In this research, this balance was determined by first collecting a large number of data points and later evaluating the best number of points after the data collection.

Surrogate models were generated using training points starting at 10 and increasing to the full 900 available data points. The other 100 remaining data points were used as validation points to calculate the fit error of the models. The number of training points increased by 10 until it reached 200, then it increased by 50 until the total number reached 900 points. This amounted to 34 surrogates with varying numbers of training points. Both the RMSE and the coefficient of determination, or R^2, were calculated for each surrogate to evaluate the accuracy of the model. The equation for R^2 is,

$$R^2 = \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2} \tag{44}$$

Table 14: Best Surrogate Settings for Each Parameter Based on RMSE

Parameter	Opt. RMSE	Opt. Correlation Function	Opt. Regression Function
C_L	0.000003	matern32	regpoly1
C_D	0.000007	exponential	regpoly1
C_m	0.001915	matern32	regpoly3
L/D	0.004057	matern52	regpoly3
P_{crit1}	5.431891	gauss	regpoly4
P_{crit2}	11.579341	linear	regpoly4
P_{crit3}	0.77634	matern52	regpoly0
P_{crit4}	13.608059	gauss	regpoly0
P_{crit5}	1.909176	matern32	regpoly3
P_{crit6}	3.844095	spline	regpoly4
P_{crit7}	36.501879	exponential	regpoly3
P_{max}	178.511649	exponential	regpoly3

where $\bar{y}$ is the average value of the validation data. R^2 is a statistic that measures how well the surrogate fits the data. A value near 1.0 indicates that the surrogate fits well with the validation data. Typically, an R^2 value of 0.99 is considered a very good fit.

The Matern-32 correlation function and 2nd order regression polynomial were specified for each of the surrogate's settings. This was determined from the analysis completed in Section 5.3. 26 different parameters were tested which included the integrated quantities (C_L, C_D, and C_m), critical surface pressures, critical surface heat flux, and critical surface heat transfer coefficients. 20 training points were used for the initial surrogate since fewer trining points resulted in failed kriging generation.

First, the integrated quantities were inspected. Figure 30 shows a log plot of the RMSE and R^2 of the surrogate model for C_L. For a small number of training points, the RMSE was large and the surrogate fit the validation data poorly. In increased number of points drove the error down significantly and the model fit improved. We can see that the R^2 value converges near 1.0 for a small number of points. The dashed black line shows a R^2 of 0.99 for reference. Only 20 points were needed to build a surrogate with an R^2 of 0.99 or greater.

The drag coefficient kriging fit metrics are shown in Figure 31. Both the RMSE and R^2 plots show large variations with changes in the number of training points. This phenomena was due to the quasi-random behavior of the sample point distribution. Adding additional training points could temporarily increase or decrease the accuracy of the surrogate depending on the spacial location and/or function value. Inspecting the R^2 plot, it is evident that the surrogate model for C_D required more training points to achieve a good fit compared to the C_L surrogate. 200 training points or more were required to ensure the drag model's R^2 was above 0.99. The RMSE plot also

Table 15: Correlation Function Rankings and Composite Scores

Parameter	Cubic	Exponential	Gauss	Linear	Matern32	Matern52	Spherical	Spline
C_L	6	3	7	8	1	5	2	4
C_D	8	1	7	4	5	2	3	6
C_m	8	6	5	4	1	2	7	3
L/D	8	7	4	5	2	1	3	6
P_{crit1}	3	8	1	7	5	2	6	4
P_{crit2}	8	7	2	1	3	6	4	5
P_{crit3}	8	7	4	5	2	1	6	3
P_{crit4}	8	7	1	6	4	3	5	2
P_{crit5}	5	3	2	7	1	4	8	6
P_{crit6}	7	2	3	8	5	6	4	1
P_{crit7}	8	1	7	4	2	3	6	5
P_{max}	8	1	2	7	6	5	3	4
Integrated	7.3	3.3	6.3	5.3	2.3	3.0	4.0	4.3
Surface	6.9	4.5	2.8	5.6	3.5	3.8	5.3	3.8
Total	7.1	4.4	3.8	5.5	3.1	3.3	4.8	4.1

shows a large increase in error from 400 to 600 training points.

To compare the surrogate performance between different parameters, the Normalized Root Mean Square Error (NRMSE) can be calculated and plotted. The NRMSE is simply the RMSE divided by the mean of the validation parameter set,

$$NRMSE = \frac{RMSE}{\bar{y}} \tag{45}$$

The C_L and C_D surrogate NRMSE trends are plotted side-by-side in Figure 32. The dashed line is plotted to show the 0.05 or 5% NRMSE location. Surrogates with a normalized error less than 5% can be considered very good models of their respective parameters of interest. Overall, the drag surrogate reported higher normalized error than the lift surrogate. This observation is likely due to nonlinearities and abnormalities within the drag distribution across the design space. The low errors in the C_L surrogate suggest that the lift is more smoothly distributed across the design space allowing for the kriging to effectively model the values.

Another method of quantitatively assessing the effectiveness of the surrogates is to directly compare the prediction values to the validation values. Scatter plots of C_L and C_D values with 20 and 200 training points are graphed in Figures 33 and 34, respectively. Each plot shows the 100 validation (actual) point values on the x-axis and the corresponding prediction on the y-axis. A dotted line is plotted to show where the predicted values should lie if they exactly equal the actual values. If a kriging surrogate with a perfect fit to the parameter was built, every prediction made

Table 16: Regression Function Rankings and Composite Scores

Parameter	regpoly0	regpoly1	regpoly2	regpoly3	regpoly4
C_L	5	1	2	4	3
C_D	3	1	2	4	5
C_m	4	2	3	1	5
L/D	5	4	2	1	3
P_{crit1}	4	5	3	2	1
P_{crit2}	5	4	3	2	1
P_{crit3}	1	3	2	5	4
P_{crit4}	1	3	2	4	5
P_{crit5}	4	5	3	1	2
P_{crit6}	3	4	2	5	1
P_{crit7}	2	3	4	1	5
P_{max}	5	3	2	1	4
Integrated	4.00	1.33	2.33	3.00	4.33
Surface	2.86	3.86	2.71	2.86	2.71
Total	3.50	3.17	2.50	2.58	3.25

using the surrogate will exactly equal the corresponding validation point values. Any modeling errors from the kriging surrogate results in over or under-prediction of the true value.

Looking at the C_L values in Figure 33, the surrogate model using 20 training points shows more scatter than the surrogate using 200 training points. This confirms that in general, surrogates using more training points results in more accurate predictions. Inspecting the predicted C_D results in Figure 34, we see the same trend of additional accuracy with more training points. However, due to the greater modeling errors in the C_D surrogates as noted above, there was a greater scattering of points compared to the C_L surrogates. This agrees with the observation from before that overall, the drag RMSE was larger than the lift RMSE.

The other integrated quantity analyzed via kriging surrogates was the pitch moment coefficient, C_m. The RMSE and R^2 for the C_m surrogate models are shown in Figure 35. Similarly to the drag coefficient, the RMSE errors for C_m displayed large variations. This indicated the potential presence of nonlinear variations within the C_m distribution due to the surrogate's sensitivity to additional training points. The R^2 plot shows a slower convergence rate than the C_L and C_D counterparts. An R^2 above 0.99 was achieved with 300 training points, but not without some variations. A converged R^2 and RMSE was observed for surrogates using 600 training points and more.

Next, the kriging surrogates for the surface pressures and heat fluxes were analyzed. Due to the

Table 17: Surrogate Model Generation Failures

Parameter	Cubic	Exponential	Gauss	Linear	Matern32	Matern52	Spherical	Spline
C_L	4	0	2	0	0	1	0	0
C_D	3	0	2	0	0	1	0	0
C_m	3	0	1	0	0	0	0	0
L/D	4	0	3	0	0	0	0	0
P_{crit1}	3	0	1	0	0	1	0	0
P_{crit2}	5	0	0	0	0	0	0	0
P_{crit3}	5	0	2	0	0	0	0	0
P_{crit4}	5	0	0	0	0	0	0	0
P_{crit5}	4	0	1	0	0	0	0	0
P_{crit6}	1	0	0	0	0	0	0	0
P_{crit7}	1	0	1	0	0	0	0	0
P_{max}	5	0	0	0	0	0	0	0

large number of output parameters, a few parameters will be chosen for inspection to provide a representation for the behavior of the surrogate models. An exhaustive collection of figures for each parameter was placed in Appendix A.

First, the upper inlet lip was examined (critical point 1). The NRMSE and R^2 for the pressure magnitude surrogate model is shown in Figure 36. This parameter was easily modeled using kriging due to the rapid reduction of error with the addition of training points. The surrogate required 70 points to achieve an NRMSE below 5% and an R^2 of above 0.99. Very little variation in the NRMSE was observed with increasing training points. The excellent performance of this kriging model is an indication of the nature of the flow at this particular point. The upper inlet lip was located at the leading edge of the entire vehicle where there were no vehicle components upstream. At this location, there were no shock-shock interactions or shock impingements due to features of the vehicle geometry. As a result, the pressure measured at this location was more predictable for a given flow condition allowing for a more accurate kriging model.

The heat flux NRMSE and R^2 at the same location were plotted in Figure 37. The primary observation to be made from these plots is that the trends mirror the pressure magnitude NRMSE and R^2 closely. Since the heat flux was calculated using the flow conditions at the surface which included the pressure magnitude, similar trends in the surrogate error are to be expected when comparing $\dot{q}$ surrogates to P surrogates.

Next, the vertical tail leading edge (critical point 6) is examined. Figure 38 shows the NRMSE and the R^2 of surrogates modeling the pressure magnitude on the tail. These plots show a slower convergence rate than the P_{crit1} surrogates. The number of training points required to reduce the NRMSE below 5% and the R^2 higher than 0.99 is 300 or more. A large number of training points

are required to accurately model the pressure on the vertical tail LE because of the increased complexity of the flow in that region. The vertical tail was located within the shock cone generated by the inlet of the GHV, and the flow within this cone was sensitive to the flow conditions and angle of attack of the vehicle. Secondary effects were thus generated around the vertical tail region due to the upstream vehicle geometry, and this induced less predictable variations in the surface pressure.

The NRMSE and R^2 for surrogates modeling the heat flux at the vertical tail LE were plotted in Figure 39. Again, the trends displayed in these two metrics were very similar to those from the pressure surrogates in Figure 38. A surrogate using 300 training points was predicted to reduce the NRMSE below 5% and R^2 above 0.99 for the heat flux on the tail.

Finally, the lower inlet throat critical point (critical point 7) was inspected using kriging. This region was characterized by very large pressure magnitudes due to the compression induced by the inlet. The NRMSE and R^2 for surrogates modeling the surface pressure were plotted in Figure 40. Both metrics show that kriging predicted the pressures at this location very poorly. The R^2 converged to a value of 0.85 and the NRMSE converged to a value of approximately 25%. The extremely poor predictive capabilities of the surrogates despite the large number of points was likely due to the high complexity of the flow in the inlet throat region. Examining the pressure contours of the symmetry plane and the internal surface in Figures 21 and 24, the lower inlet throat critical point was located near an extremely strong shock formed from the lower lip of the inlet. The point was also near the expansion region at the inlet to isolator transition. Significant back-pressure within the isolator could have also played a role in influencing the pressure at this point. Overall, there are multiple, strong effects that influence the pressure at this critical point that are nonlinear and unpredictable with respect to the flow conditions.

The heat flux surrogate NRMSE and R^2 values for different numbers of training points are plotted in Figure 41. Interestingly, the kriging surrogates predicted the heat flux excellently at the lower inlet throat location. Only 150 training points were needed to provide enough accuracy in the surrogate model for a 5% or lower NRMSE and 0.99 or higher R^2.

Each output parameter surrogate models were inspected to determine how many training points were required to achieve good accuracy. Table 18 shows each of these parameters and their respective required training point numbers. The NRMSE and R^2 plots for each parameter not shown in this section are provided in Appendix A in Figures A.7 to A.20.

From this study we see that some parameters were more easily predicted using kriging compared to others. For integrated quantities, the lift was easiest to predict, and only 20 points were needed to effectively model C_L across the design space. The drag and the pitch moment coefficients were more difficult to predict, and the surrogates needed around 300 training points for overall

Table 18: Number of Training Points Required to Meet Accuracy Standard

Parameter	Points Required
C_L	20
C_D	190
C_m	600
L/D	100
P_{crit1}	70
P_{crit2}	300
P_{crit3}	40
P_{crit4}	400
P_{crit5}	80
P_{crit6}	300
P_{crit7}	N/A
$\dot{q}_{crit1}$	130
$\dot{q}_{crit2}$	300
$\dot{q}_{crit3}$	50
$\dot{q}_{crit4}$	180
$\dot{q}_{crit5}$	60
$\dot{q}_{crit6}$	300
$\dot{q}_{crit7}$	130

adequate model accuracy. The exception to this generality was the pitch moment coefficient which required over 600 points to achieve accuracy. The surrogates modeling the surface pressures were better at predicting points with fewer complexities in the surrounding flowfield. The number of training points needed for an accurate surface pressure surrogate varied depending on the location. The inlet leading edge pressure surrogate only required 50 training points while the inlet throat pressure surrogate could not achieve accurate predictions for any number of training points. All in all, the number of CFD training points required for an accurate surrogate interpolation is completely dependent on the parameter of interest.

It is important to note that the number of required training points for surrogate model accuracy is dependent on the fidelity of the CFD code used. As the CFD fidelity increases, additional features are resolved that add complexity to the outputs. For instance, if a low-fidelity panel method was used, certain features such as shock waves and compressibility effects are not resolved and the resulting CFD outputs will have a simpler, more linear response. Thus fewer training points are required for a lower-fidelity CFD code. The opposite is true as well. In the detailed vehicle design phase, a higher fidelity model is required which means that more training points will be needed to

build an adequately accurate surrogate.

5.5 Surrogate Inspection

This final study will quantitively examine the behavior of the output parameters using kriging surrogates. Each surrogate model shown below used all 900 training points for maximum prediction accuracy, a Matern-32 correlation function, and a 2nd order regression polynomial. Since the design space consisted of three design variables: Mach, angle of attack, and altitude; visualizing the distribution of values of a parameter is difficult. Two approaches were taken to display the results of the kriging surrogates that were generated: contours and isosurface plots. The contours will plot the parameter on a Mach x-axis and α y-axis with constant altitudes. The iso surfaces will show surfaces of constant parameter values within the three dimensional design space.

First, the coefficient of lift was examined. Contours of C_L were plotted in Figure 42. As expected, the lift was shown to be linearly dependent on α. However, the slope of C_L to α was steeper with a low Mach number and shallower with a higher Mach number. In other words, changes in α resulted in a stronger change in C_L for a low Mach number, and a weaker change in C_L for a large Mach number. The contours plotted at each altitude were unchanged meaning that C_L was independent of altitude. A white dotted line was plotted on the contours to show where $C_L = 0.0$. The zero lift point occurred around an angle of attack of $-2°$. This makes sense because the incidence of the GHV wing was 2.4° which resulted in lift generation at a 0° angle of attack. The isosurfaces shown in Figure 43 clearly show that C_L was independent of altitude and that the largest values of C_L occurred at high α and low Mach.

Next, the drag coefficient was plotted as contours in Figure 44 and isosurfaces in Figure 45. C_D was shown to be highly dependent on both α and Mach number. C_D had a parabolic trend with respect to angle of attack with larger α resulting in more drag. The Mach number affected the slope of the parabolic trend, lower Mach numbers caused higher C_D to α dependency, and higher Mach numbers caused lower C_D to α dependency. Again the drag coefficient was shown to be independent of altitude with the exception of some minor variations in the contour shapes.

The lift to drag ratio, L/D, was modeled via kriging and displayed in Figure 46 and Figure 47. The lift to drag ratio had a more unique distribution within the design space. The largest values were observed at low altitudes at a high α and high Mach. L/D was highly dependent on α with less of a dependency on Mach and altitude. At higher altitudes, between 19 and 25 km, a prominent "ridge" of high L/D (~ 7.0) was observed around $\alpha = 2.5°$ and a Mach number between 6.0 and 7.5. This suggests that the GHV can cruise efficiently at these conditions. The conditions corresponding to $L/D = 0$ was shown via the white dotted line. The location of

$L/D = 0$ was the same as the location of $C_L = 0$.

The final integrated quantity, C_m, was shown in Figure 48 and Figure 49. C_m was significantly more non-linear than C_L and C_D which was evident from the wavy patterns in the contours. The distribution of C_m is similar to the C_L distributions, but with an increased sensitivity to the Mach number. The pitch moment coefficient was shown to be independent of altitude. The relationship between C_m and α was nearly linear for a given Mach number. Lower Mach numbers resulted in strong dependency of C_m on α, and higher Mach numbers resulted in much lower dependency of C_m on α. Higher α caused lower, negative C_m values which corresponded to a *pitch down* moment. The vehicle has a pitch up response for positive values of C_m which were caused by negative α. The white line indicates the trim condition ($C_m = 0$) of the GHV which was located near -2.3° for low Mach and lower for higher Mach. Note that the C_m values reported here were calculated using a center of mass located at the geometric centroid of the GHV. In reality, the center of mass can be adjusted to provide the desired stability characteristics.

One of the most significant parameters that describes the static stability of an aircraft is C_{m_α}, or the pitch stability derivative with respect to angle of attack. Throughout the design space, C_{m_α} was shown to be negative. In other words, as α increased, C_m decreased which indicated that the GHV was statically stable in the pitch axis [37]. For example, if the GHV was flying at trim condition and it experienced a positive perturbation in α, the negative C_{m_α} will cause a negative C_m reaction which will pitch down the aircraft nose back to trim. The most important observation from this analysis is that the magnitude of C_{m_α} decreases with increased Mach number, which means that the GHV was less statically stable at higher speeds.

Additional contour plots for each critical pressure and heat flux are shown in Figures A.37 to A.50, and isosurface plots are shown in Figures A.51 and A.52.

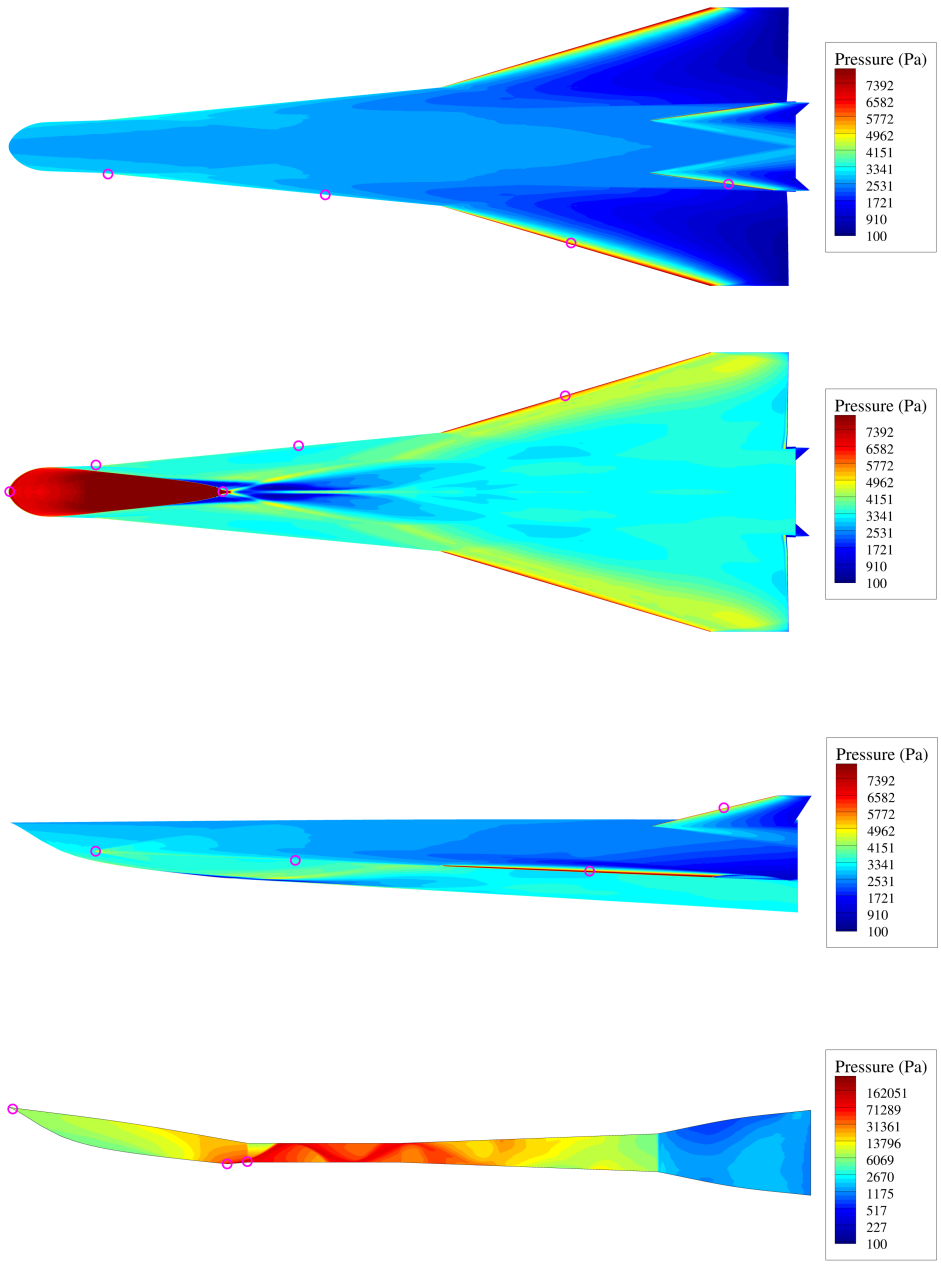


Figure 28: Surface Pressure Contours (top, bottom, side, internal), $M = 6.0$, $\alpha = 0.0°$, $h = 25$ km

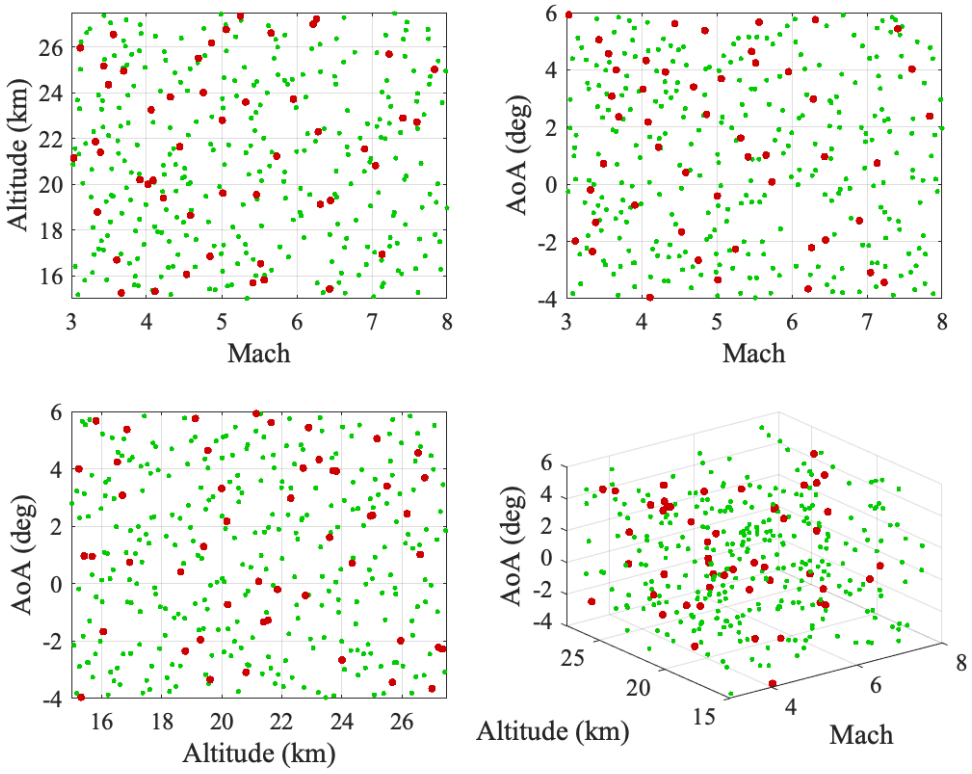
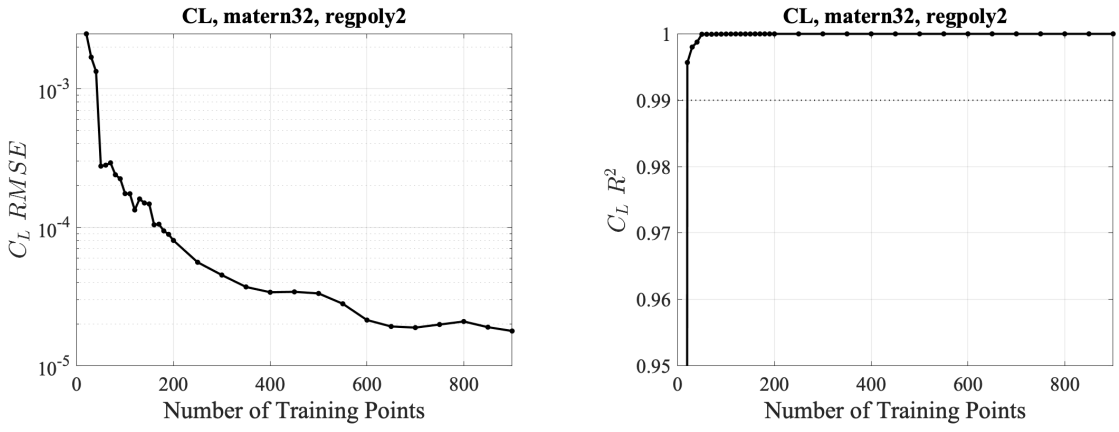
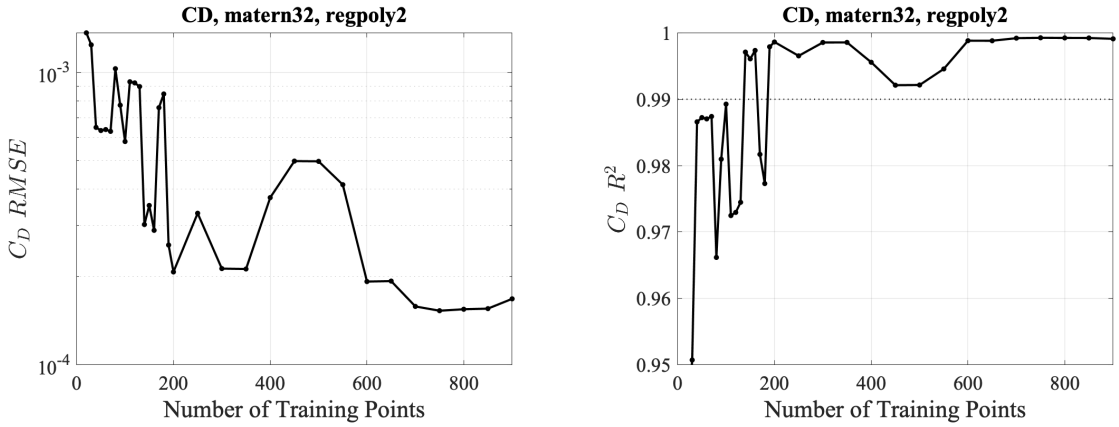


Figure 29: Plot of 300 Training Points (green) and 50 Validation Points (red)



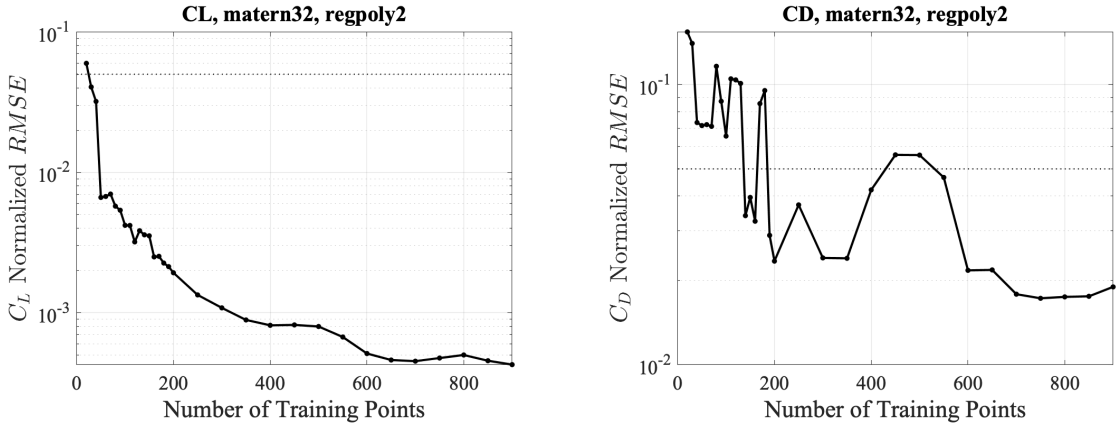
(a) RMSE (b) R^2

Figure 30: Coefficient of Lift



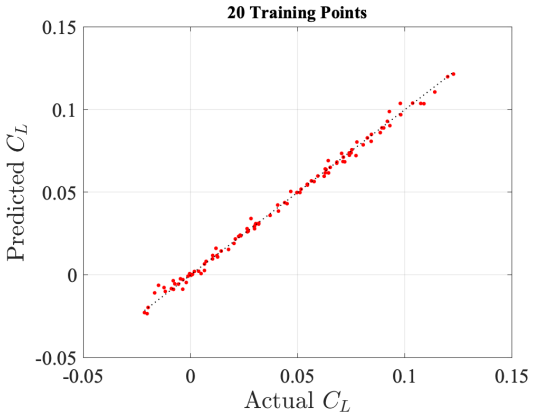
(a) RMSE (b) R^2

Figure 31: Coefficient of Drag

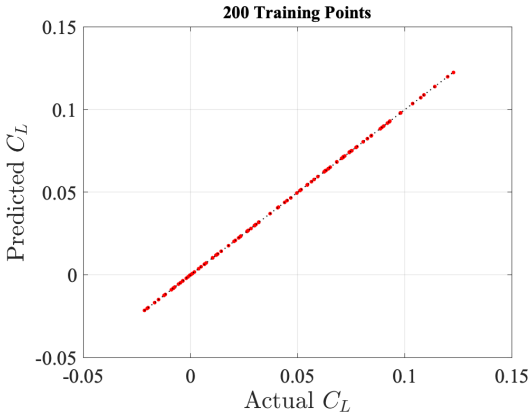


(a) C_L (b) C_D

Figure 32: Surrogate NRMSE for C_D and C_L

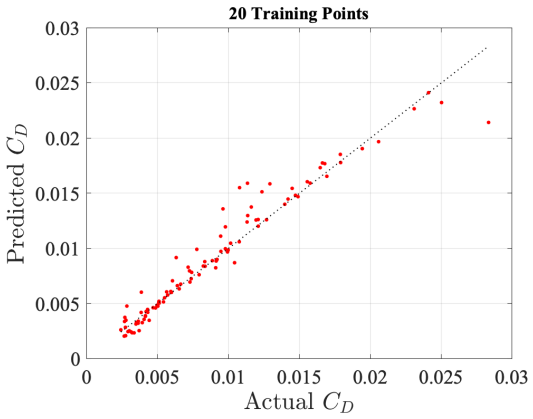


(a) 20 Training Points

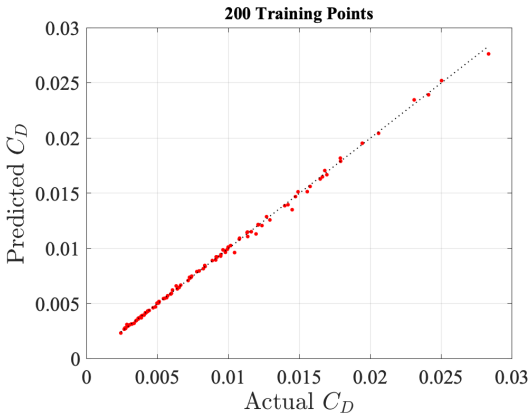


(b) 200 Training Points

Figure 33: C_L Surrogate Predictions using 20 and 200 Training Points

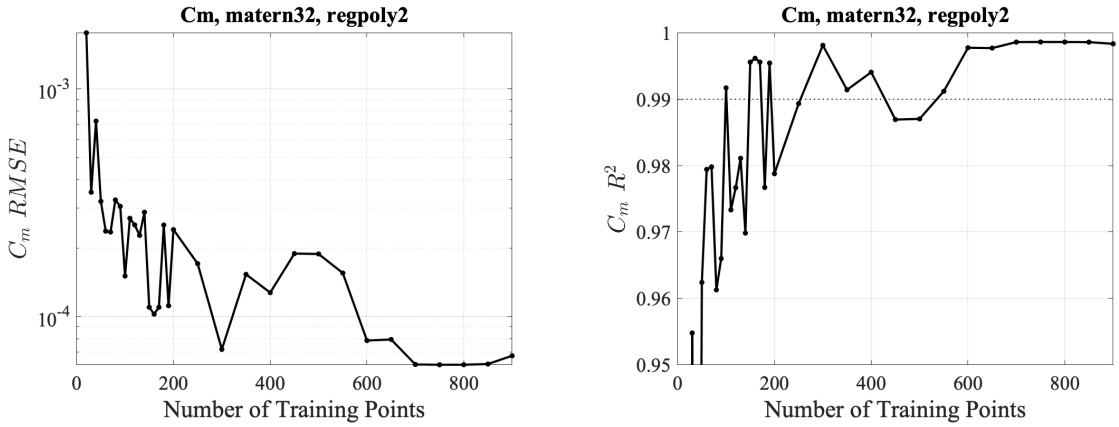


(a) 20 Training Points



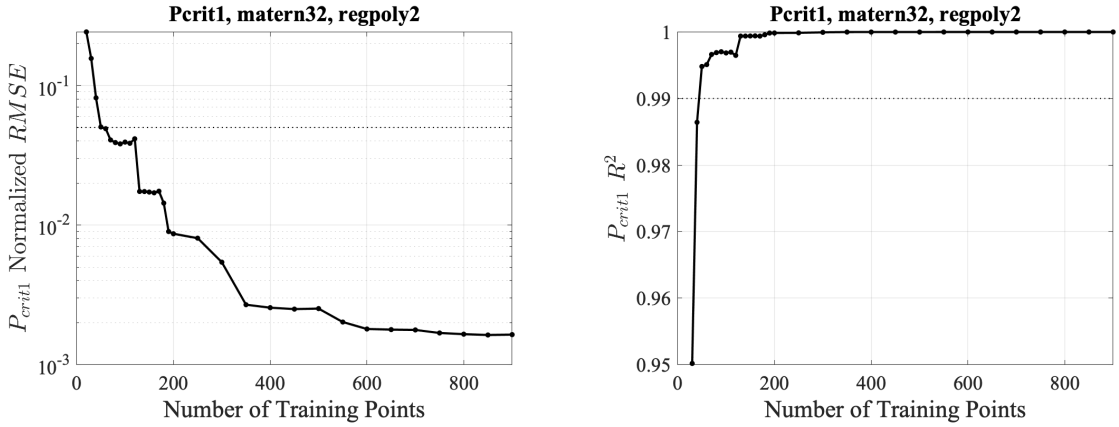
(b) 200 Training Points

Figure 34: C_D Surrogate Predictions using 20 and 200 Training Points



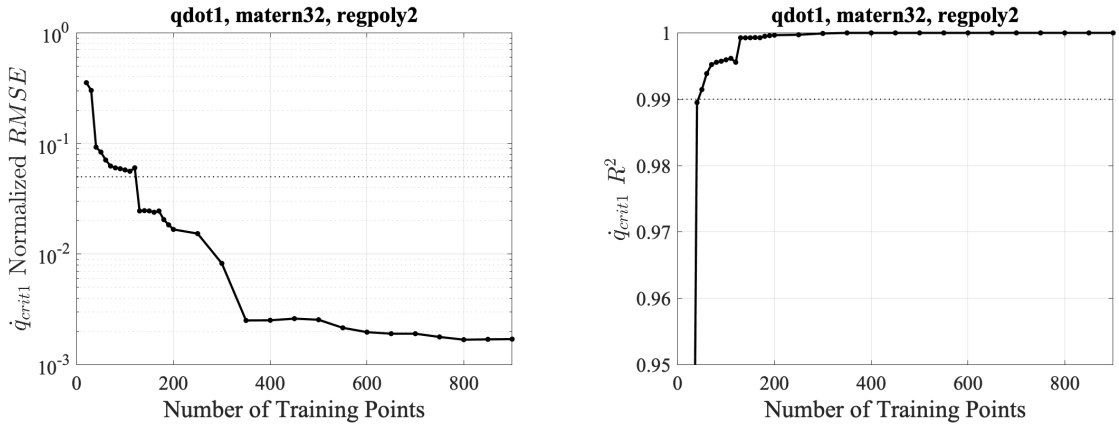
(a) RMSE (b) R^2

Figure 35: Pitch Moment Coefficient



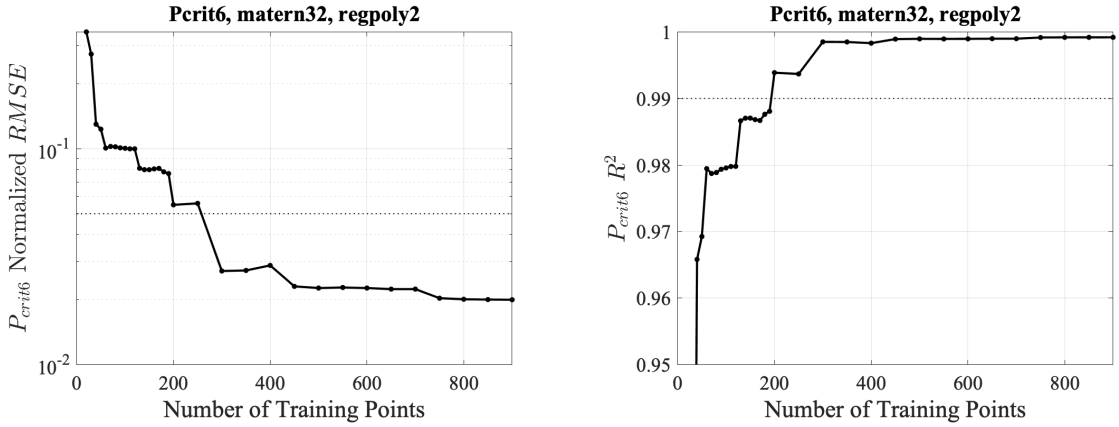
(a) RMSE (b) R^2

Figure 36: Upper Inlet Lip Pressure



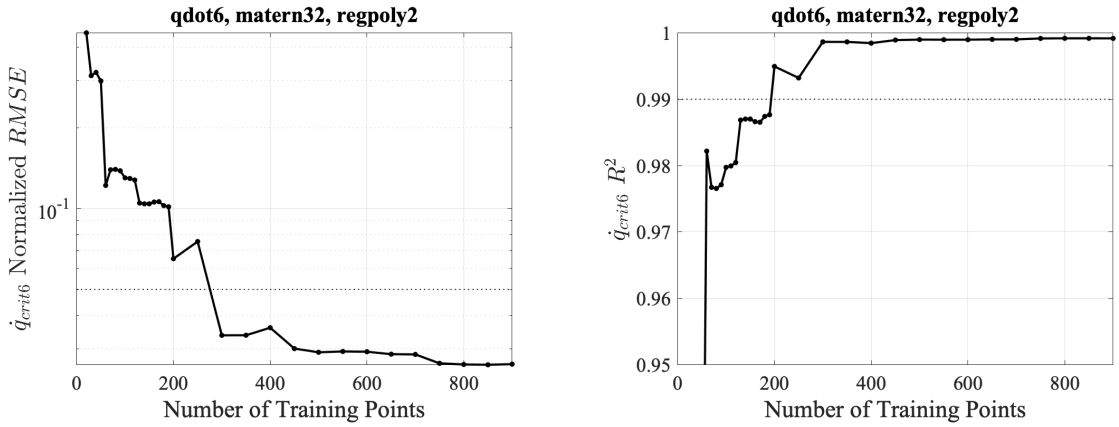
(a) RMSE (b) R^2

Figure 37: Upper Inlet Lip Heat Flux



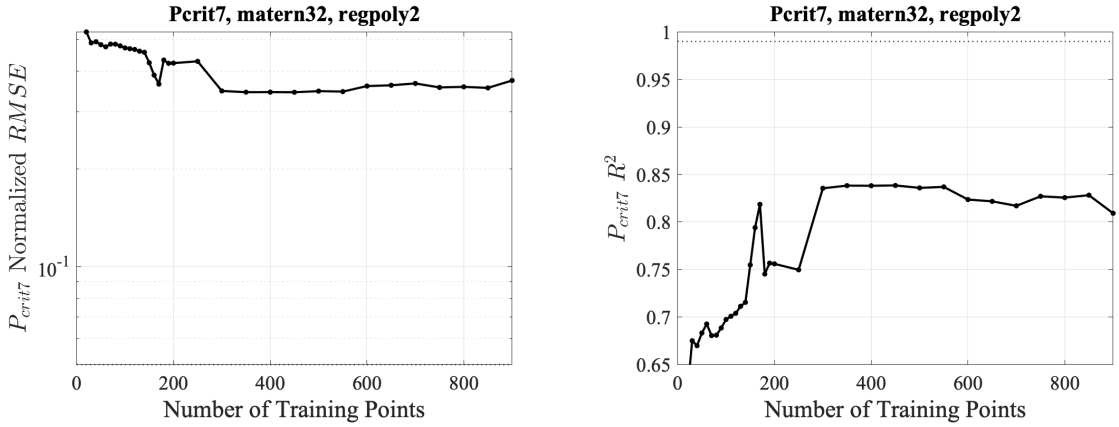
(a) RMSE (b) R^2

Figure 38: Vertical Tail LE Pressure



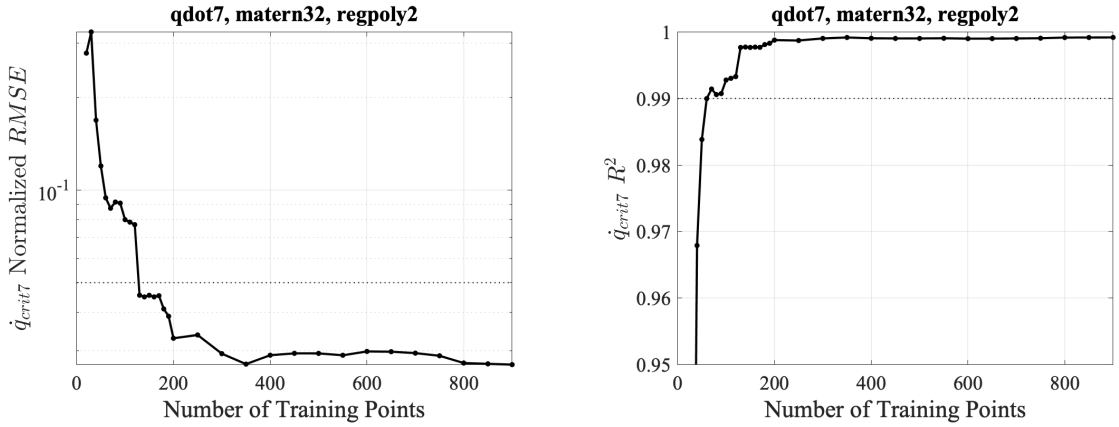
(a) RMSE (b) R^2

Figure 39: Vertical Tail LE Heat Flux



(a) RMSE (b) R^2

Figure 40: Lower Inlet Throat Pressure



(a) RMSE (b) R^2

Figure 41: Lower Inlet Throat Heat Flux

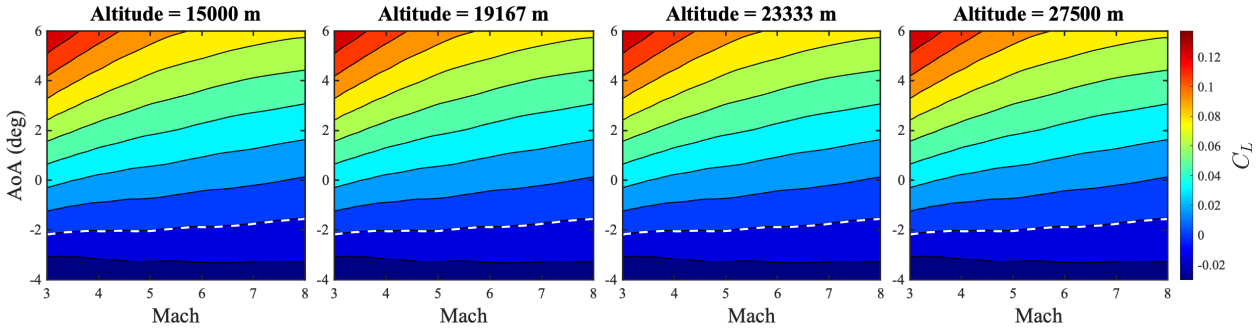


Figure 42: C_L Surrogate Contours

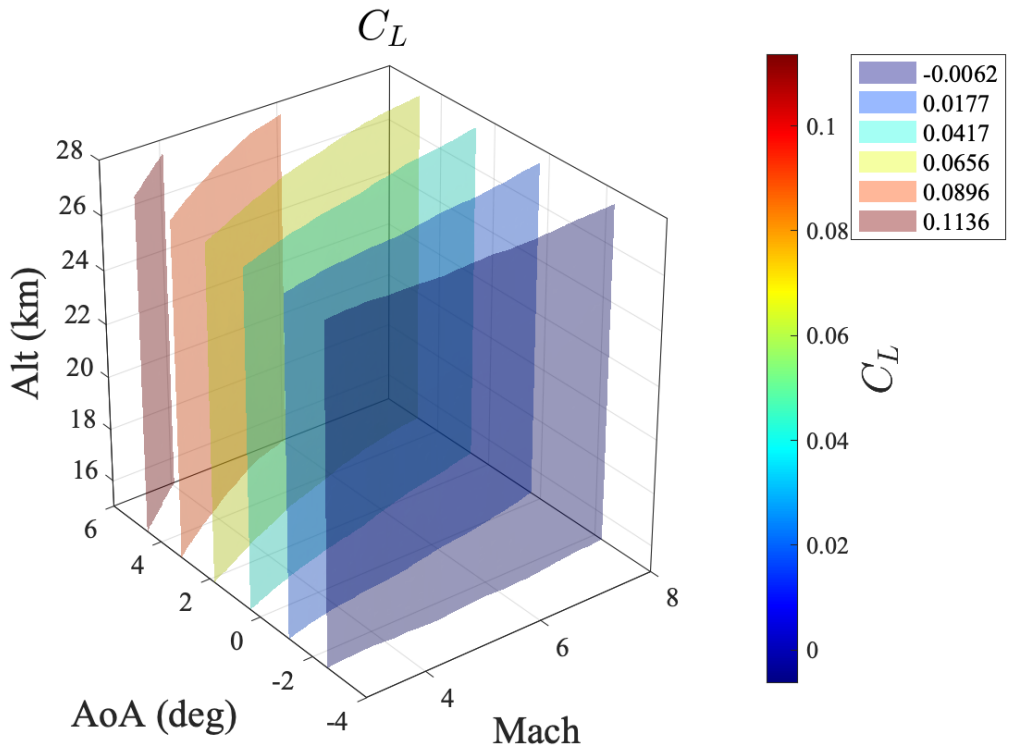


Figure 43: C_L Surrogate Isosurfaces

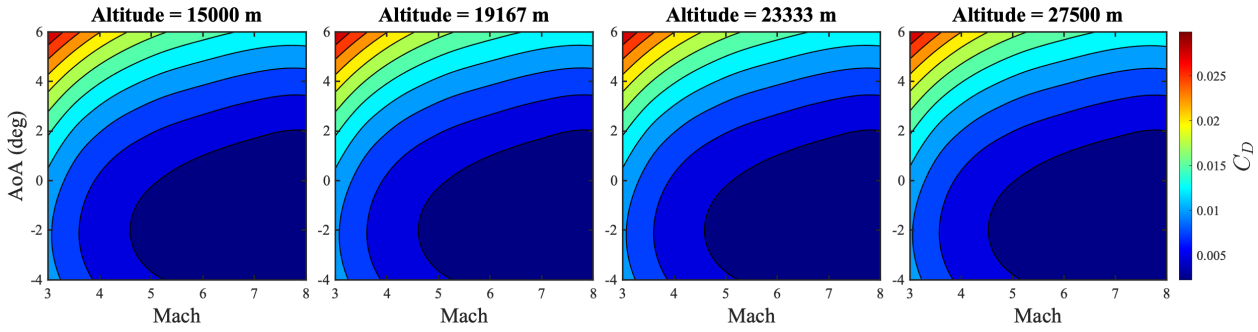


Figure 44: C_D Surrogate Contours

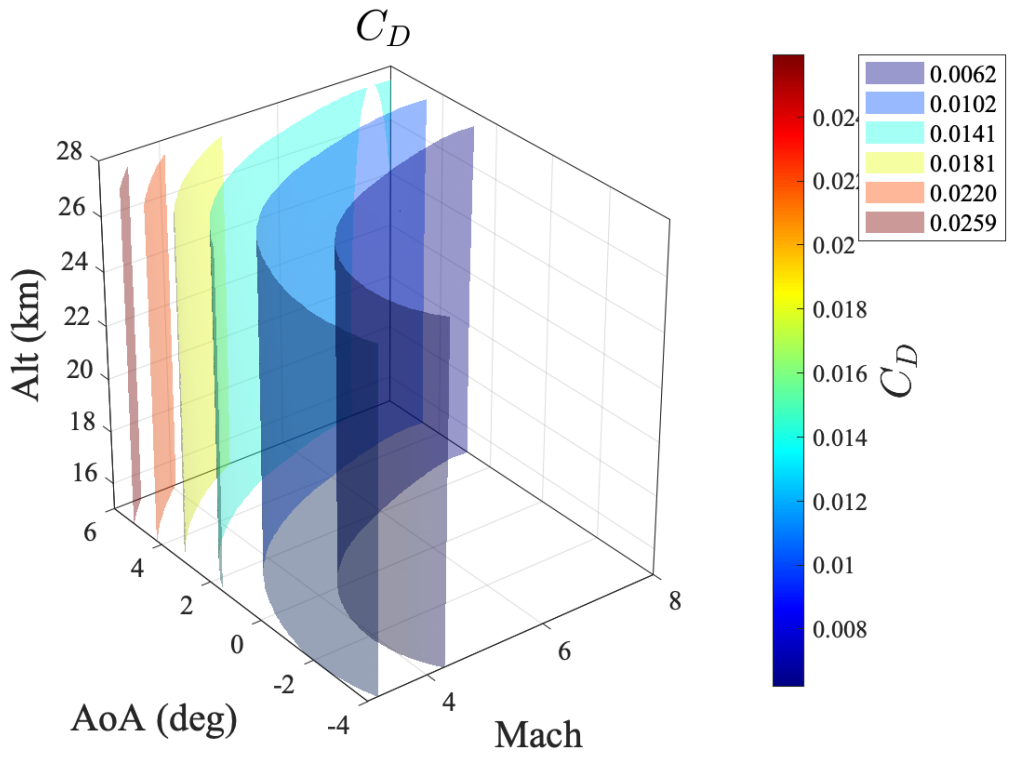


Figure 45: C_D Surrogate Isosurfaces

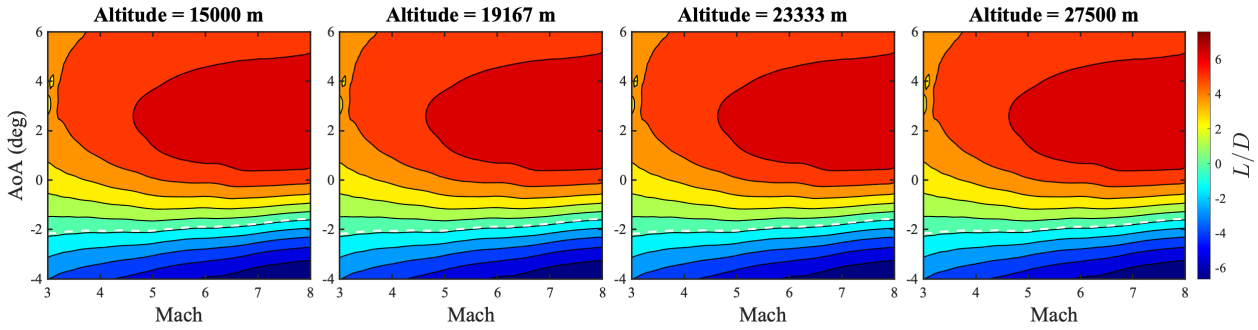


Figure 46: L/D Surrogate Contours

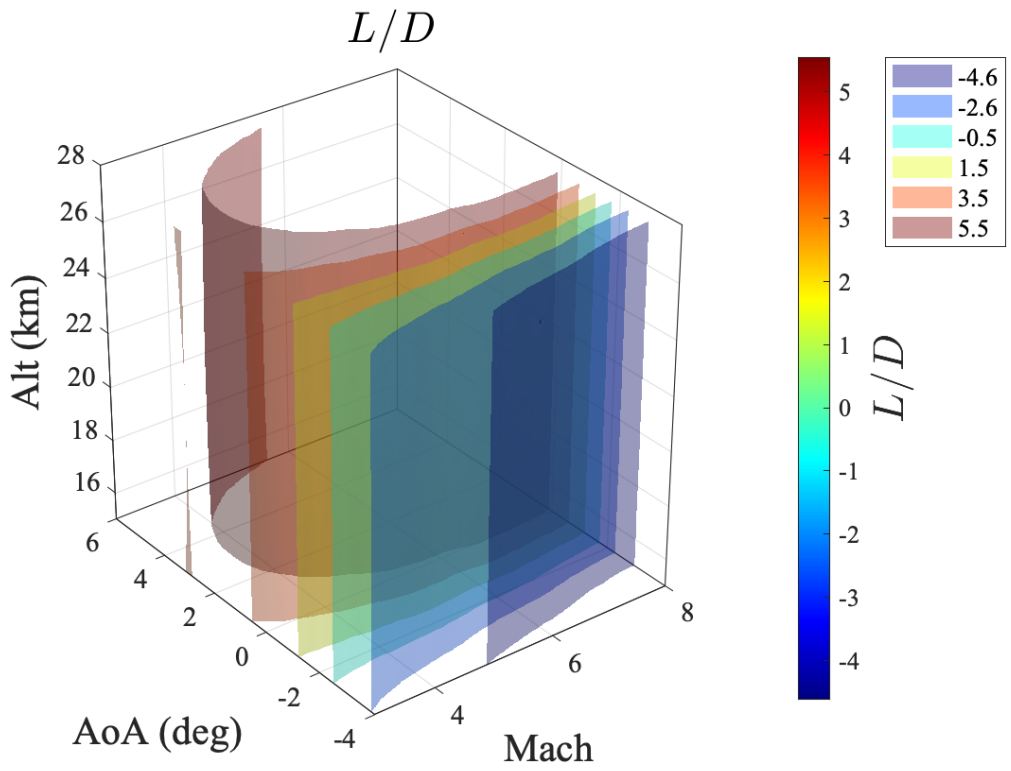


Figure 47: L/D Surrogate Isosurfaces

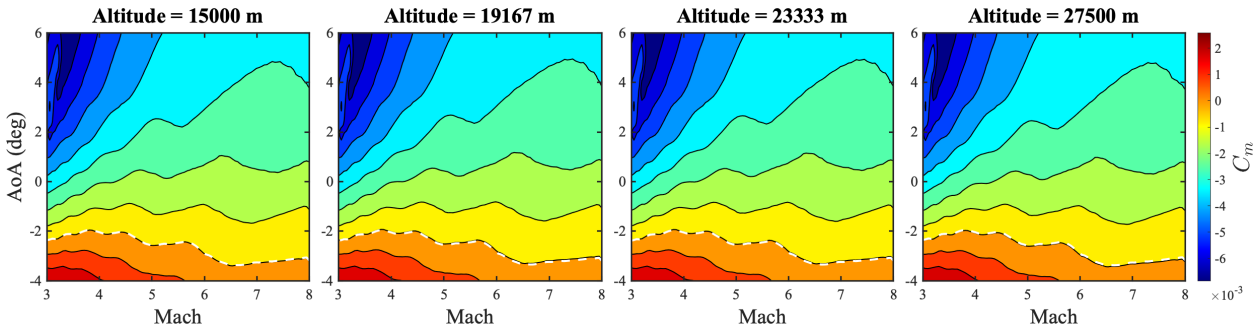


Figure 48: C_m Surrogate Contours

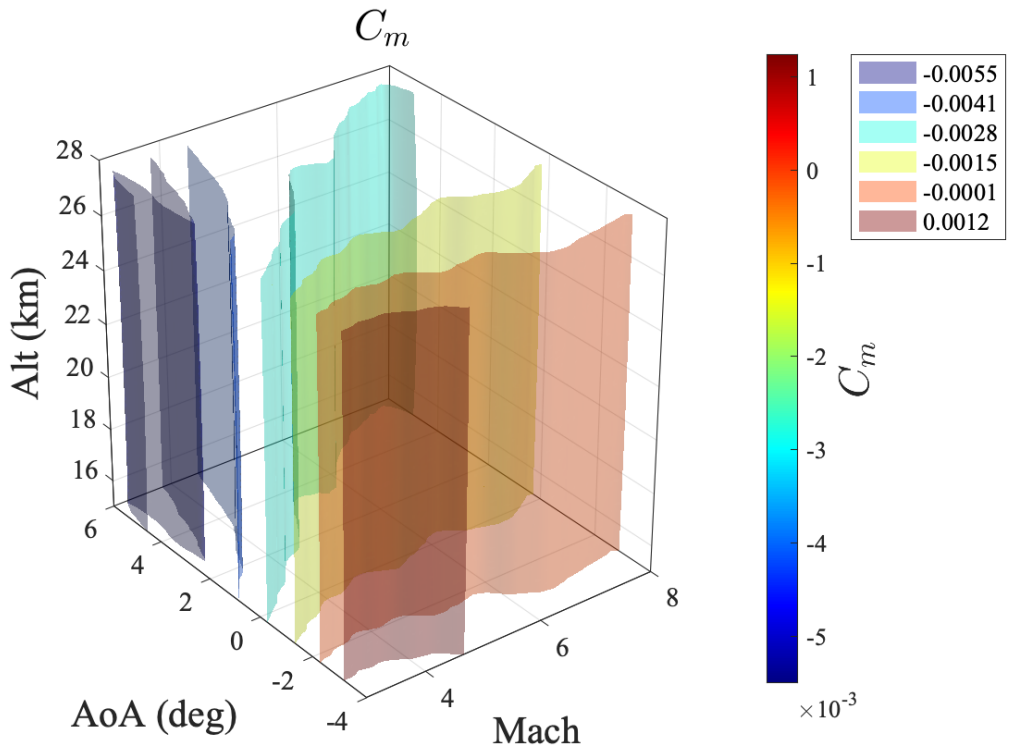


Figure 49: C_m Surrogate Isosurfaces

6 CONCLUSIONS

This research demonstrated the process of modeling the aerodynamic responses, specifically integrated quantities such as C_D and C_L and surface conditions, of a GHV across a range of flight conditions. The CFD process was demonstrated which included the grid development, grid convergence study, and flow solution. Additionally, the methodology behind the selection of the optimal space-filling set of sample points was shown. Kriging surrogates with differing settings were run to determine the optimal combination of correlation function and regression function for each parameter of interest. Using this information, a study was conducted that assessed the quality of the kriging surrogates for a varying number of training points. Finally, the kriging surrogates were inspected and the aerodynamic responses from the GHV were analyzed.

The portion of the study that required the most time was the grid generation. In 1991, CFD expert Joe Steger stated that "grid generation remains the key pacing item in making CFD useful to most engineering applications in aerodynamics" [38], and this statement was very true during the conduction of this research. Approximately 3 months of work was required to develop an initial grid that modeled the 3-D GHV and fully conduct the grid convergence study. During the grid convergence study, the trend of the C_L, C_D, and surface pressure with respect to the grid refinement contained variations that prevented a monotonic convergence. This observation was most likely a result of the behavior of the algorithms used to develop the unstructured volumetric flowfield.

The qualitative analysis of the pressure flowfield revealed features of the flow that influenced the performance of the GHV. It was observed that the scramjet inside the GHV operated as expected in the unstarted condition. The inlet successfully compressed the oncoming flow via isentropic compression, the isolator maintained the high internal pressure via a shock train, and the nozzle allowed for the expansion and acceleration of the flow. At the nominal flow conditions, the GHV behaved like a wave-rider by using the strong shock from the inlet and the incidence of the wing to generate a high pressure region underneath the vehicle. Seven points of high pressures were identified based on the surface distribution across the GHV surface. Six out of the seven points were located on the leading edge of a feature (wing, inlet, tail).

After collecting the CFD data across the design space and implementing the kriging models, several takeaways were observed. It was found that for the integrated quantities, C_L, C_D, and C_m, the optimal correlation function was the Matern-32 correlation function using a 1st order polynomial as regression function. For surface quantities, it was found that the exponential correlation function combined with a 2nd or 4th order regression polynomial was optimal. Finally, if all the responses needed to be modeled using common settings, the Matern-32 function combined with a 2nd order regression was the best option. It was found that some correlation functions performed over an order of magnitude better than others, while some correlation functions performed very similarly to each other. Thus it was shown that it is important to conduct an analysis on different settings to ensure that the kriging surrogates are sufficiently accurate.

The sizing study on kriging surrogates showed how many training points were needed to build an adequately accurate model for predictions. It was determined that a surrogate with a normalized

RMSE of 5% or lower and an R^2 of 0.99 or higher was considered converged. With this standard in mind, the C_L required the least amount of training points of only 20. The C_D surrogate needed approximately 200 points and the C_m surrogate needed approximately 300 training points for a converged surrogate. The surrogates for surface pressures converged using fewer points if the surrounding flow complexity was low, and they converged slower if the flow complexity was high. The inlet LE pressure surrogate required 50 points while the tail LE pressure surrogate required 300 training points for convergence. The inlet throat pressure surrogate was an extreme case where the surrogate was not able to converge to an NRMSE below 5% and R^2 above 0.99. This was due to the very complex flow interactions that occurred near that critical point.

Finally, the kriging models were inspected and observations were made on the aerodynamic responses. It was found that for each of the coefficients (C_L, C_D, C_m), there was no dependence on altitude. The lift, drag, and pitching moment were highly dependent on the angle of attack. However, as the Mach number increased, the dependency of the coefficients on the angle of attack decreased. Additionally, the static stability of the GHV decreased with increasing Mach number from inspection of C_m. After examining the lift to drag ratio, it was found that there was an area of large L/D approximately equal to 7.0 around the flight condition $\alpha = 2.5°$, a Mach number between 6.0 and 7.5, and an altitude between 19 and 25 km.

6.1 Future Work

Due to the multidisciplinary nature of hypersonic vehicle design, there are a myriad of different ways the method of kriging can be applied to modeling a vehicle like the GHV. Some areas of research that can expand the understanding of preliminary design of hypersonic vehicles and progress the work done in this research are listed below.

- The analysis that was conducted in this research used Euler CFD code which left out important effects from viscosity. The GHV can be modeled using RANS and kriging can be used to build surrogates. The RANS and Euler surrogates can be compared and conclusions drawn from their differences. The major difficulty with using RANS is the task of converging the grid and procuring enough CPU-Hours to run enough data points.
- The GHV can be modeled with a started scramjet via combustion physics. The propulsion aspect of the GHV can be analyzed which can provide insight into the performance of the engine throughout a range of operating conditions. Design variables such as fuel equivalence ratio, fuel injection rate, fuel temperature, and flameholder measurements can be added and the thrust, combustion pressure ratio, exit velocity and other parameters can be modeled via kriging.
- The static and dynamic stability discipline of the GHV can be described via kriging. The symmetry plane can be removed and the grid can be modified to include elevons and rudders to the GHV design. The sideslip, AoA, and control surface deflections can be used as design variables for a kriging surrogate, and the set of stability responses (C_L, C_m, C_n, etc.) can be predicted.

- Finally, the controllability of the GHV can be defined using surrogate models. By using kriging to describe the aerodynamic responses, stability, propulsion, and other disciplines pertaining to the GHV, a control loop can be developed. A state vector containing flow conditions, control surface deflections, scramjet controls, and other parameters can be developed and the corresponding GHV response can be modeled using the kriging function. Optimal control theory can then be utilized to optimally plan a particular mission with certain requirements.

REFERENCES

[1] Fidan, B. and Mirmirani, M. and Ioannou, P. Flight Dynamics and Control of Air-Breathing Hypersonic Vehicles: Review and New Directions. Number 7081 in AIAA Paper 2003-7081, Norfolk, Virginia, 2003. AIAA.

[2] Z.D. Witeof and C.L. Pasiliao. Fluid-Thermal-Structural Interaction Effects in Preliminary Design of High Speed Vehicles. AIAA Paper 2015-1631, Kissimmee, Florida, 2015. AIAA.

[3] J.D. Anderson. *Hypersonic and High-Temperature Gas Dynamics*. AIAA Education Series. AIAA, 3rd edition, 2019.

[4] B. Glaz, L. Liu, and P.P. Friedman. A Surrogate Based Approach to Reduced-Order Dynamic Stall Modeling. AIAA Paper 2010-3042, Orlando, Florida, 2010. AIAA.

[5] W.C. Williams. Mercury Project Summary. Technical report, NASA, Washington D.C., 1963. SP-45.

[6] Y. Gibbs. NASA Armstrong Fact Sheet: First Generation X-1. https://www.nasa.gov/centers/armstrong/news/FactSheets/FS-085-DFRC.html, 2014.

[7] R. Calzada. NASA - Dryden History - Historic Aircraft - X-1 Technical Data. https://www.nasa.gov/centers/dryden/history/HistoricAircraft/X-1/techdata.html, 2008.

[8] NASA. X-15, Research at the Edge of Space. 1963. EP-9.

[9] C. Peebles. *Eleven Seconds into the Unknown: A History of the Hyper-X Program*. Library of Flight. AIAA, Reston, Virginia, 2011.

[10] W.H. Heiser and D.T. Pratt. *Hypersonic Airbreathing Propulsion*. AIAA Education Series. AIAA, Wright-Patterson Air Force Base, Ohio, 1994.

[11] J.D. Anderson. A Survey of Modern Research in Hypersonic Aerodynamics. AIAA Paper 1984-1578, Snowmass,Colorado, 1984. AIAA.

[12] R.M. Cummings, W.H. Mason, S.A. Morton, and D.R. McDaniel. *Applied Computational Aerodynamics*. Cambridge Aerospace Series. Cambridge University Press, 2015.

[13] D.J. Kinney. Aerothermal Anchoring of CBAERO Using High Fidelity CFD. *NASA Ames Research Center*, 2007.

[14] N.J. Greiner. *Convective Heat Transfer with and without Film Cooling in High Temperature, Fuel Rich and Lean Environments*. PhD thesis, Air Force Institute of Technology, 2014.

[15] E.R.G. Eckert. Engineering Relations for Heat Transfer and Friction in High-Velocity Laminar and Turbulent Boundary-Layer Flow Over Surfaces with Constant Pressure and Temperature. *Transactions of the American Society of Mechanical Engineers*, 78(6), 1956.

[16] E.R.G. Eckert. Survey of Boundary Layer Heat Transfer at High Velocities and High Temperatures. *Wright Air Development Center*, 1960.

[17] N.J. Falkiewicz. *Reduced-Order Aerothermoelastic Analysis of Hypersonic Vehicle Structures*. PhD thesis, The University of Michigan, 2012.

[18] S. Fitzgerald. Calculation of Air Properties. https://github.com/sjfitz/AirProperties, 2017.

[19] N. Cressie. The Origins of Kriging. *Mathematical Geology*, 22(3), 1990.

[20] E.B. Anderes. Kriging. *Encyclopedia of Environmetrics*, 2013.

[21] A.I.J. Forrester, A. Sóbester, and A.J. Keane. *Engineering Design via Surrogate Modelling*. Wiley, University of Southhampton, UK, 2008.

[22] I. Couckuyt, T. Dhaene, and P. Demeester. ooDACE Toolbox: A Matlab Kriging toolbox: Getting started. 2013.

[23] D.G. Krige. A Statistical Approach to Some Basic Mine Valuation Problems on the Witwaterstrand. *Journal of the Chemical Metallurgical and Mining Society of South Africa*, 52(3), 1951.

[24] G. Matheron. Principles of Geostatistics. *Economic Geology*, 58, 1963.

[25] E.R. Dreyer. Rapid Steady-State Hypersonic Aerothermodynamic Loads Prediction Using Reduced Fidelity Models. *Journal of Aircraft*, 58(3), 2020.

[26] S. Powell and A. Sóbester. *Computational Intelligence in Aerospace Sciences*. AIAA, 2014.

[27] P. Bratley and B.L. Fox. ALGORITHM 659: Implementing Sobol's Quasirandom Sequence Generator. 1988.

[28] B. Ruttle, J. Stork, and G. Liston. Generic Hypersonic Vehicles for Conceptual Design Analyses. Technical report, Air Force Research Laboratory, Wright-Patterson Air Force Base, OH, 2012.

[29] F. Ferguson, N. Dasque, H. Mrema, and J.P. Kizito. A Coupled Aerodynamic and Propulsive Performance Analysis of a Generic Hypersonic Vehicle. AIAA Paper 2015-3839, Orlando, Florida, 2015. AIAA.

[30] F. Ferguson, N. Dasque, M. Dhanasar, and L. Uitenham. An Aerodynamic Analysis of the Generic Hypersonic Vehicle. AIAA Paper 2018-0637, Kissimmee, Florida, 2018. AIAA.

[31] D.G. Sagerman, N. Dasque, M.P. Rumpfkeil, and B. Hellman. Comparisons of Measured and Modeled Aero-thermal Distributions for Complex Hypersonic Configurations. AIAA Paper 2017-0264, Grapevine, Texas, 2017. AIAA.

[32] J. Heley, L. Gonzalez, and B. Chudoba. Generic Hypersonic Vehicle Design Configuration Verification. AIAA Paper 2018-5258, Orlando, Florida, 2018. AIAA.

[33] I.M. Boyd, R.V. Grandhi, and J. Camberos. Generic High-Speed Vehicle Configuration Modeling and Optimization. AIAA Paper 2020-3153, 2020.

[34] U.S. Standard Atmosphere, 1976. Technical memorandum, NASA, Washigton, D.C., 1976. TM-X-74335.

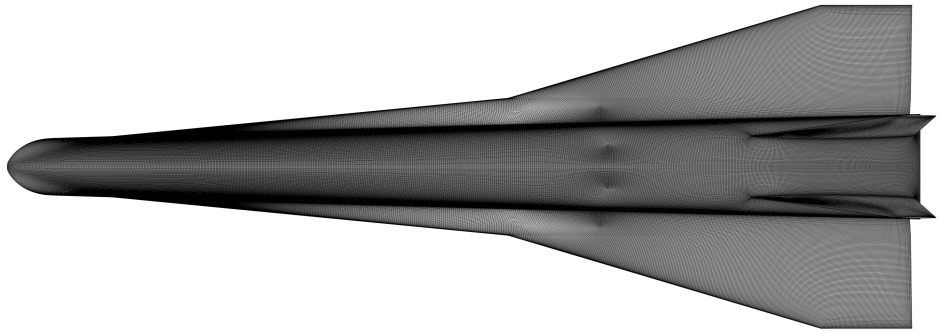
[35] R. Courant, K.O. Friedrichs, and H. Lewy. Über die Partiellen Differenzengleichungen der Mathematischen Physik. *Mathematische Annalen*, 100, 1928.

[36] M. Graczyk, T. Lasota, and B. Trawinski. Comparative Analysis of Premises Valuation Models Using KEEL, RapidMiner, and WEKA. *Wrocław University of Technology*, 2009.

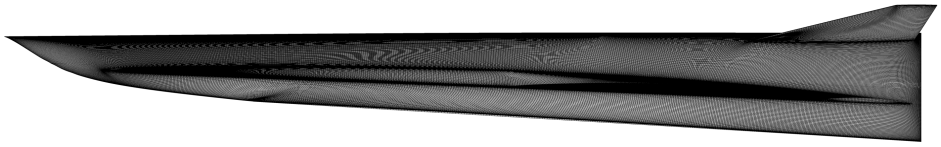
[37] T.R. Yechout. *Introduction to Aircraft Flight Mechanics*. AIAA Education Series, 2014.

[38] J.L. Steger. Technical Evaluation Report: AGARD Fluid Dynamics Panel Specialty Meeting on Applications of Mesh Generation to Complex3-D Configurations. *AGARD-AR-268*, 1991.

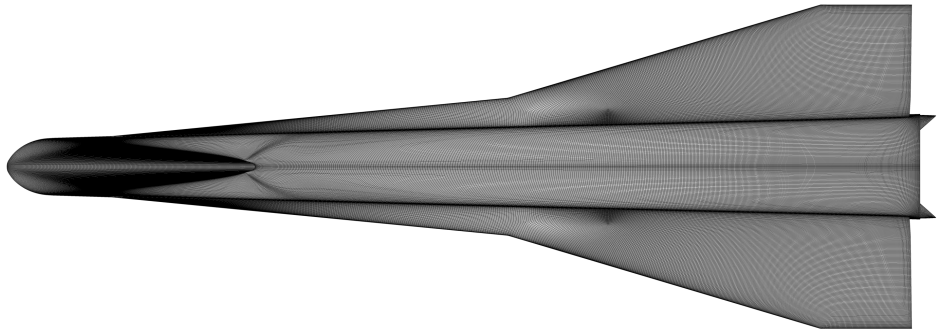
APPENDIX A: REFERENCED FIGURES



(a) Top View



(b) Side View



(c) Bottom View

(d) Internal View

Figure A.1: GHV Surface Grid (Not Coarsened)

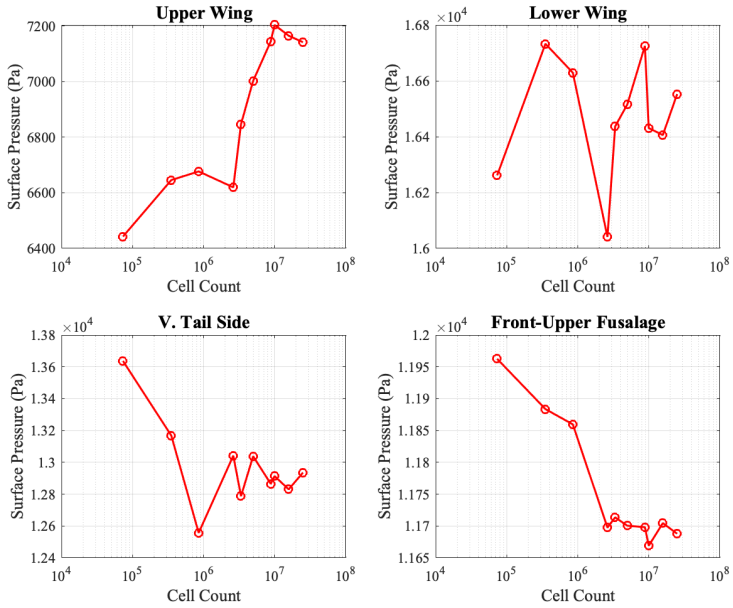


Figure A.2: Euler Grid Convergence of Surface Pressures for TC-2

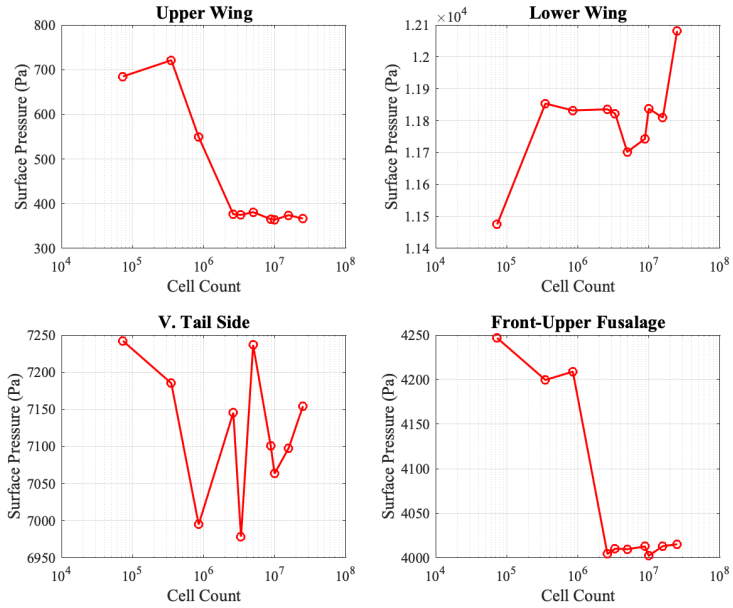


Figure A.3: Euler Grid Convergence of Surface Pressures for TC-3

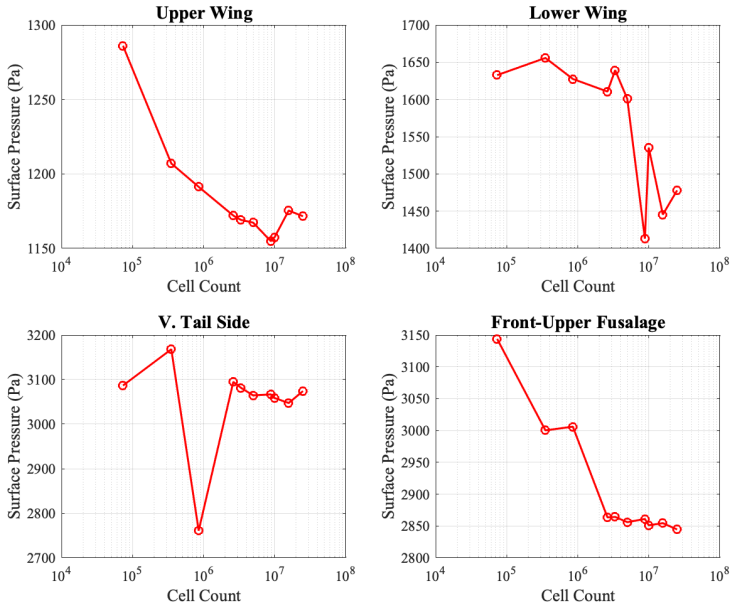


Figure A.4: Euler Grid Convergence of Surface Pressures for TC-4

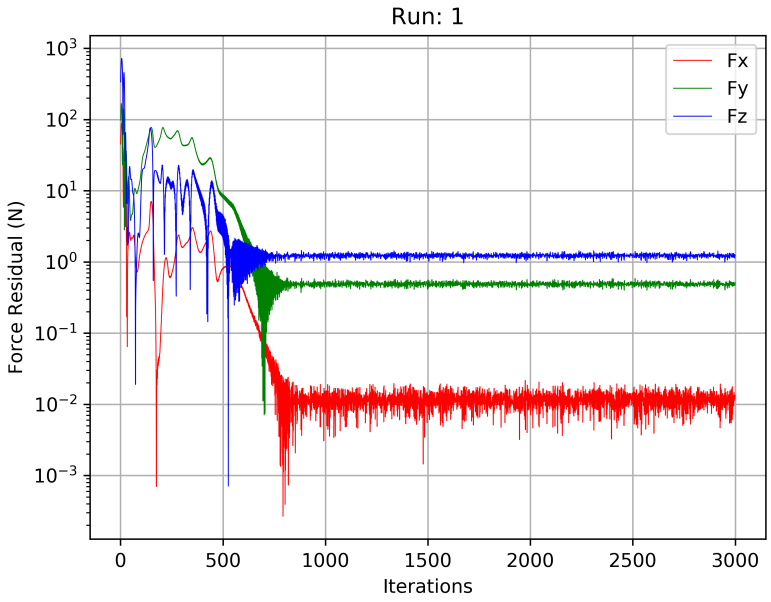
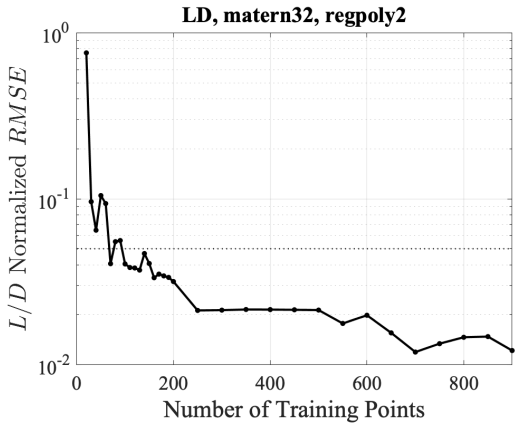
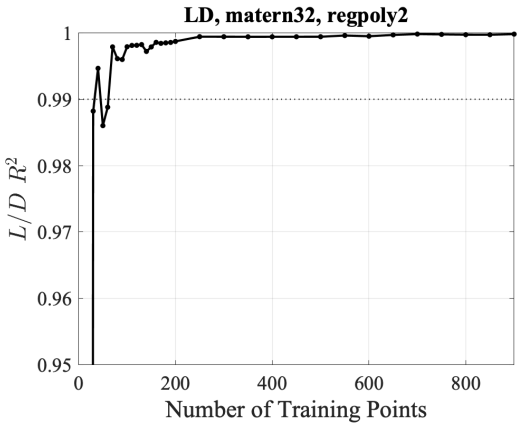


Figure A.5: RUN-0001 Force Residual Plot

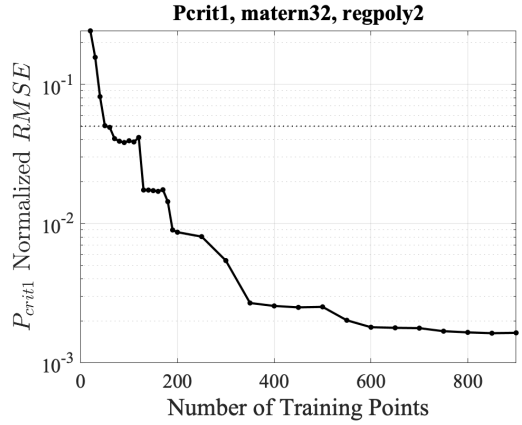


(a) NRMSE

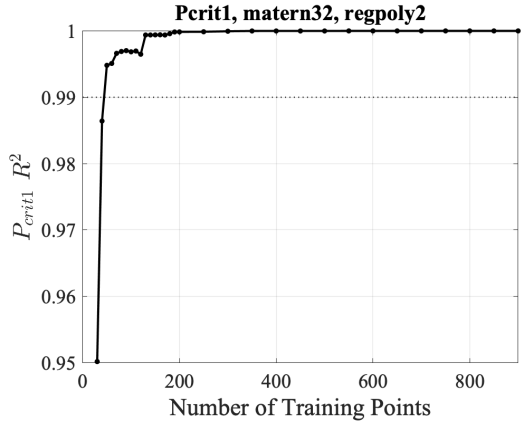


(b) R^2

Figure A.6: Lift to Drag Ratio

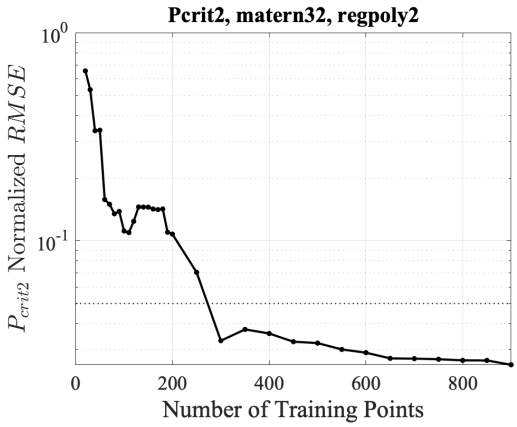


(a) NRMSE

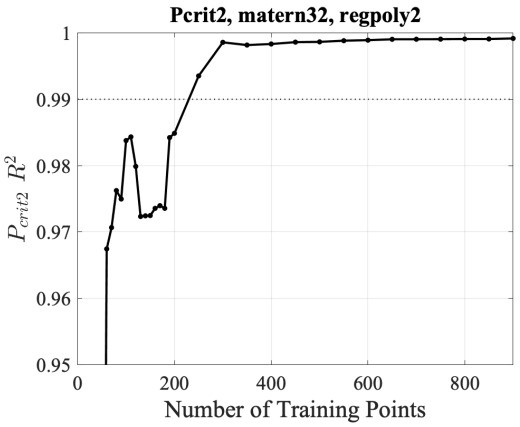


(b) R^2

Figure A.7: Upper Inlet Lip Pressure

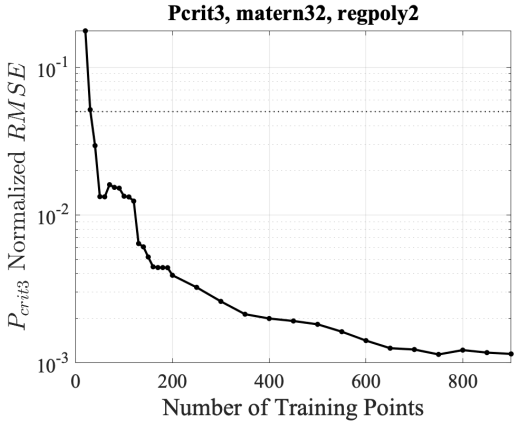


(a) NRMSE

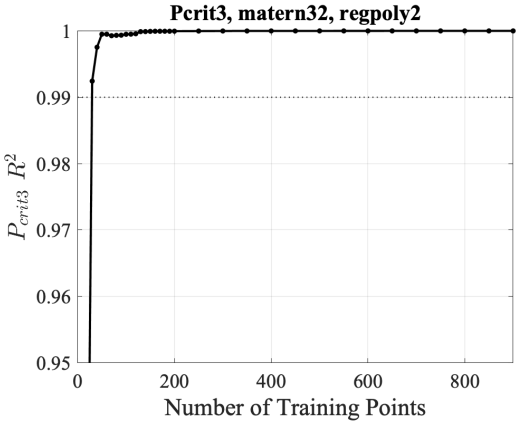


(b) R^2

Figure A.8: Lower Inlet Lip Pressure

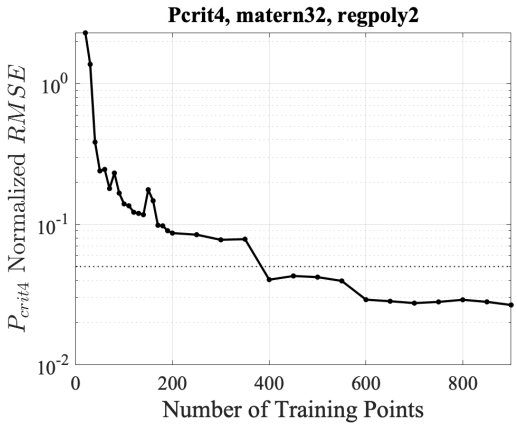


(a) NRMSE

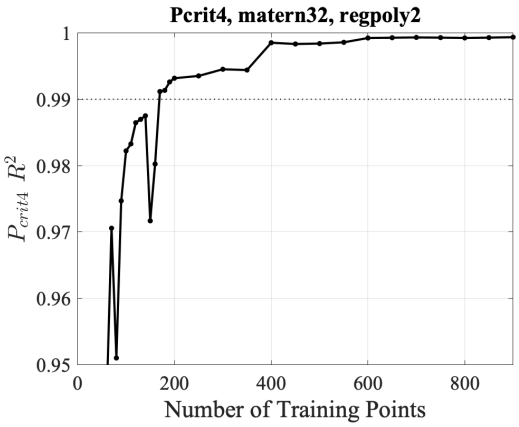


(b) R^2

Figure A.9: Fuselage/Strake Junction Pressure

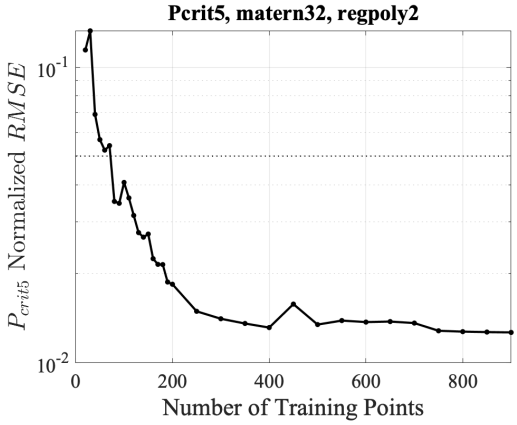


(a) NRMSE

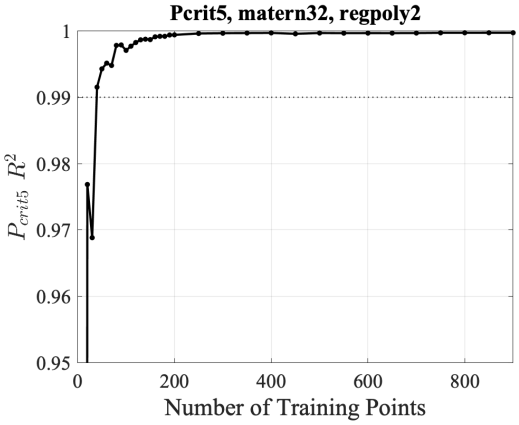


(b) R^2

Figure A.10: Strake LE Pressure

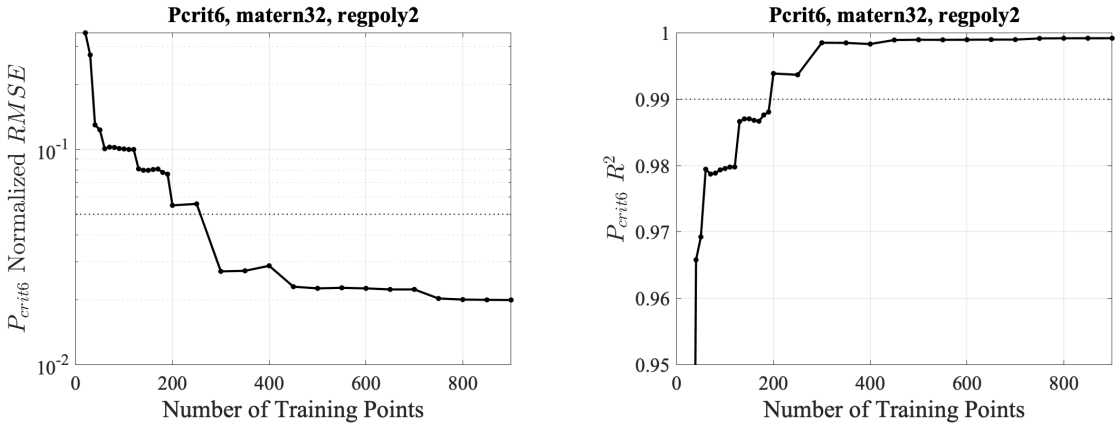


(a) NRMSE



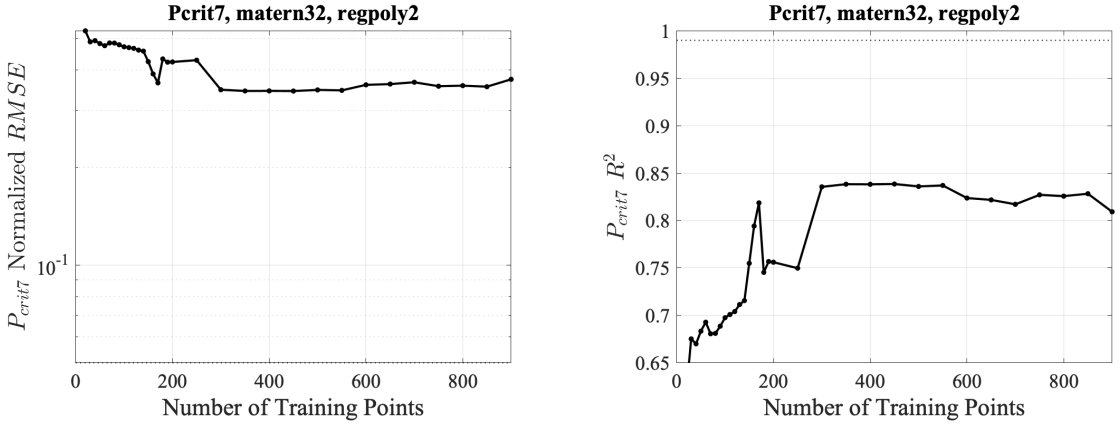
(b) R^2

Figure A.11: Wing LE Pressure



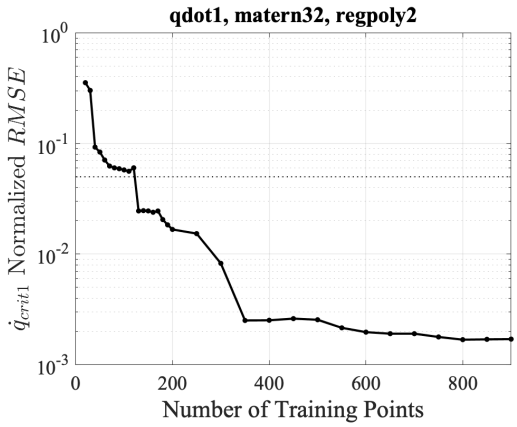
(a) NRMSE (b) R^2

Figure A.12: Vertical Tail LE Pressure

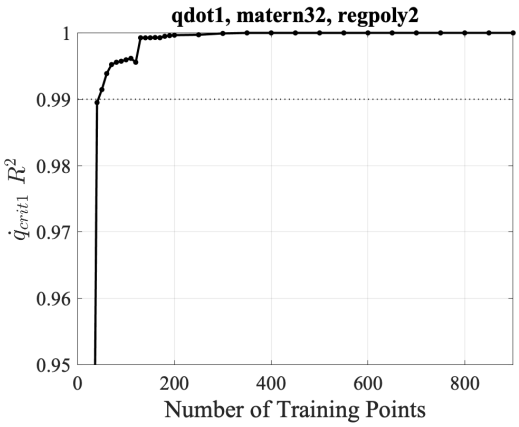


(a) NRMSE (b) R^2

Figure A.13: Lower Inlet Throat Pressure

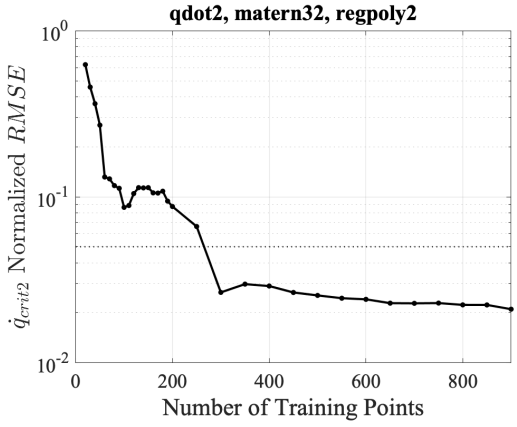


(a) NRMSE

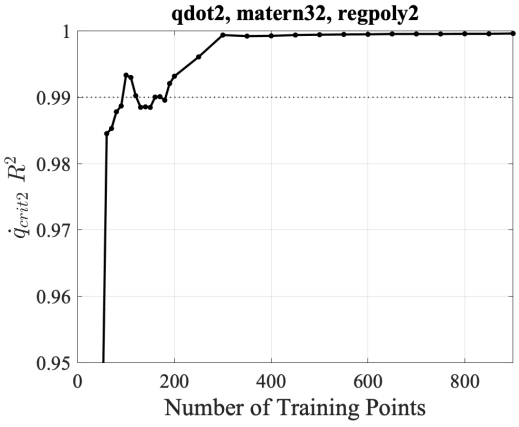


(b) R^2

Figure A.14: Upper Inlet Lip Heat Flux

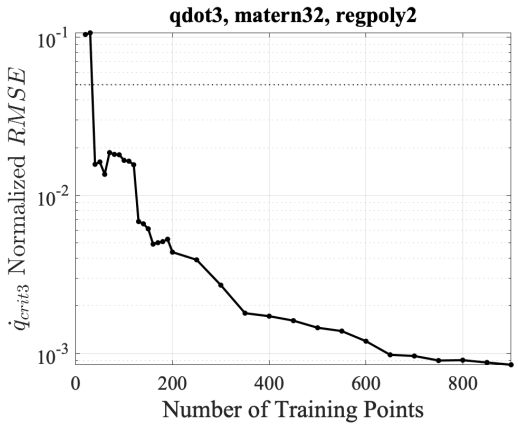


(a) NRMSE

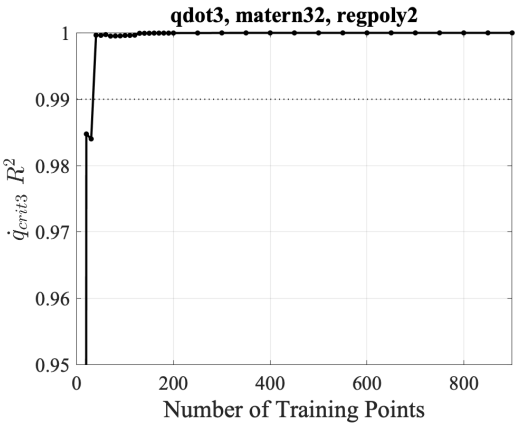


(b) R^2

Figure A.15: Lower Inlet Lip Heat Flux

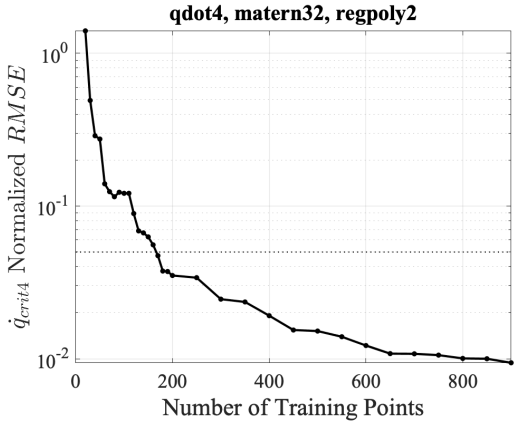


(a) NRMSE

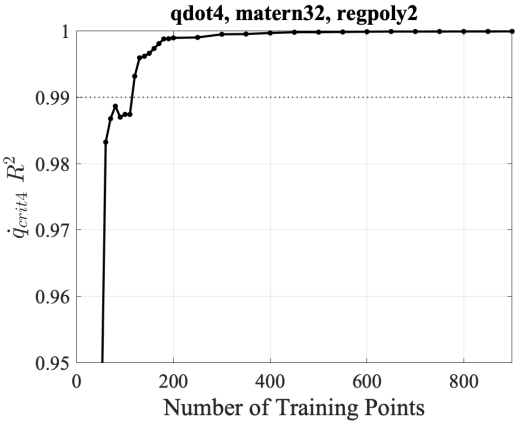


(b) R^2

Figure A.16: Fuselage/Strake Junction Heat Flux

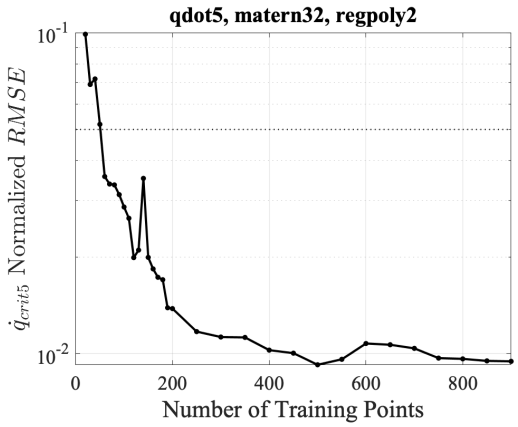


(a) NRMSE

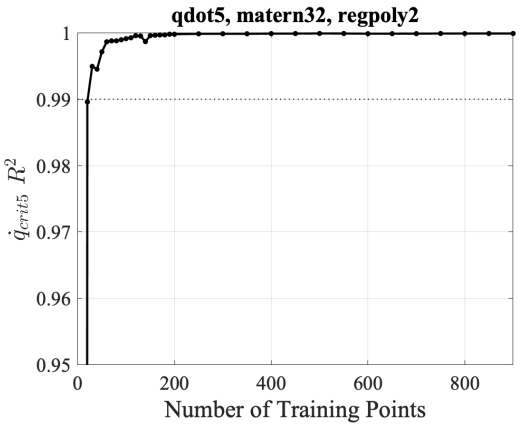


(b) R^2

Figure A.17: Strake LE Heat Flux

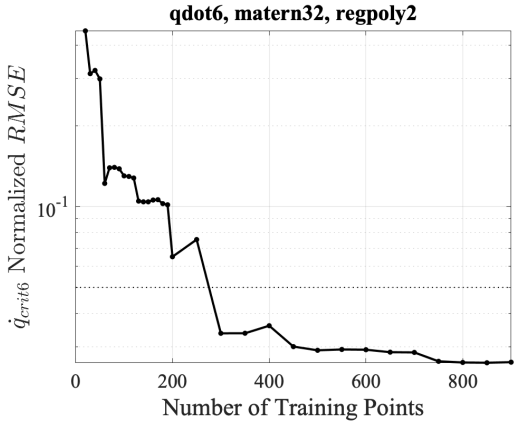


(a) NRMSE

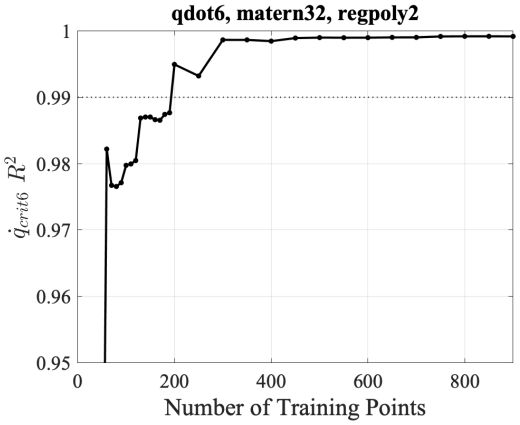


(b) R^2

Figure A.18: Wing LE Heat Flux

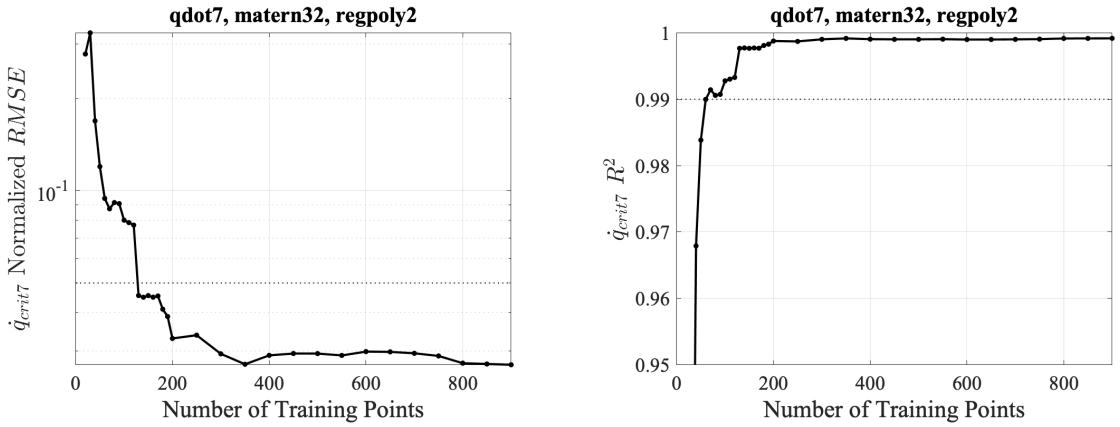


(a) NRMSE



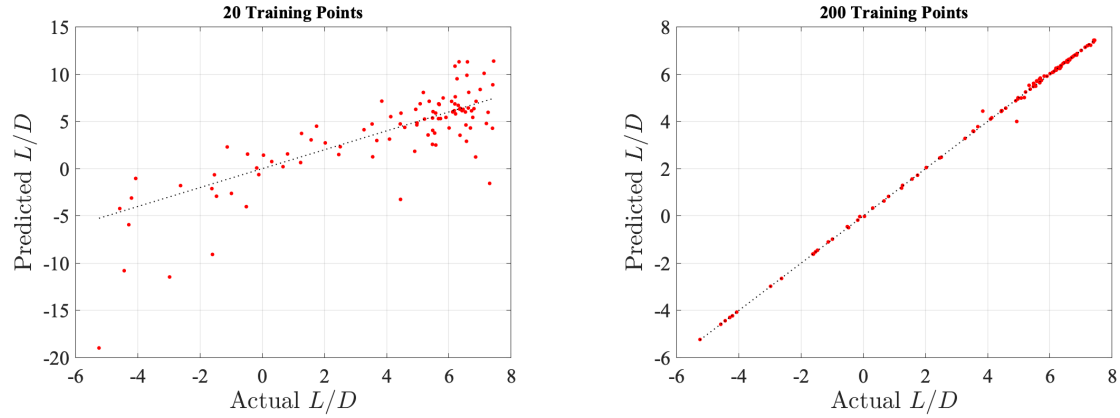
(b) R^2

Figure A.19: Vertical Tail LE Heat Flux



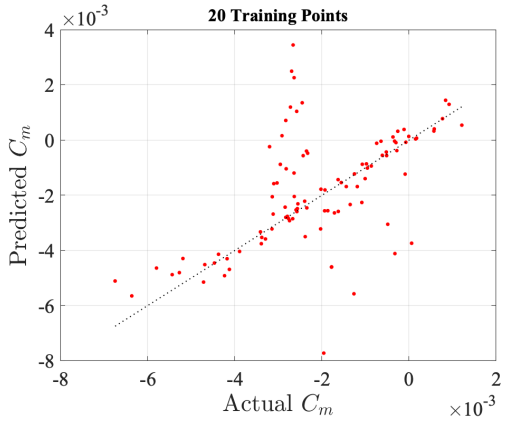
(a) NRMSE (b) R^2

Figure A.20: Lower Inlet Throat Heat Flux

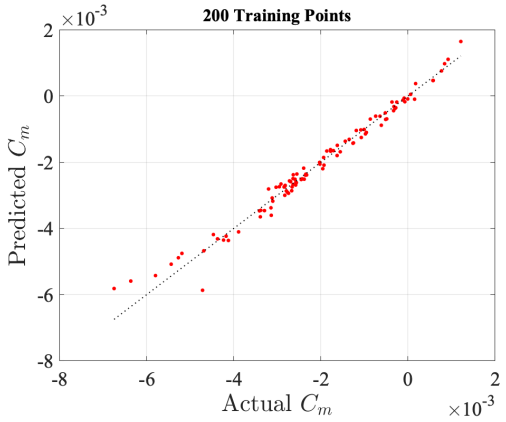


(a) 20 Training Points (b) 200 Training Points

Figure A.21: L/D Surrogate Predictions

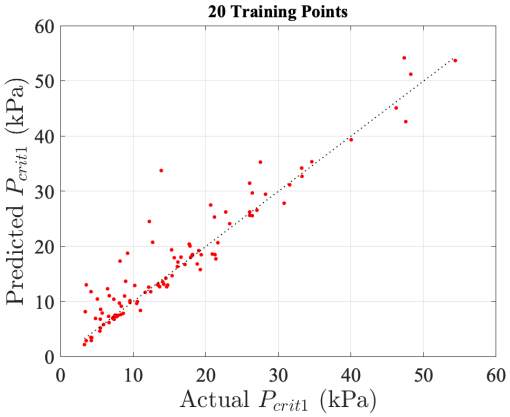


(a) 20 Training Points

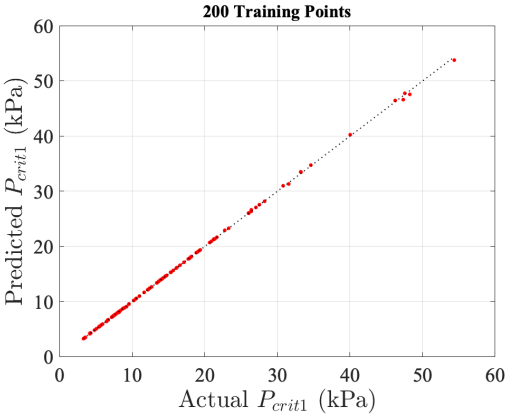


(b) 200 Training Points

Figure A.22: C_m Surrogate Predictions

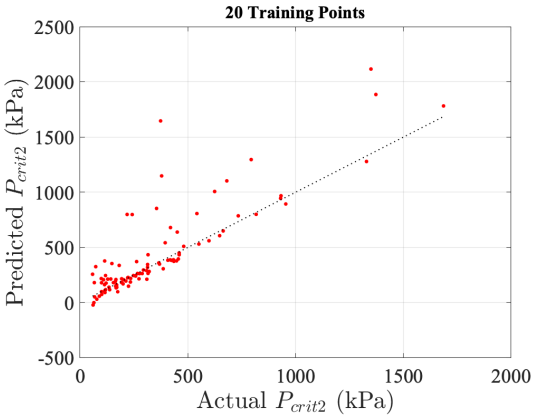


(a) 20 Training Points

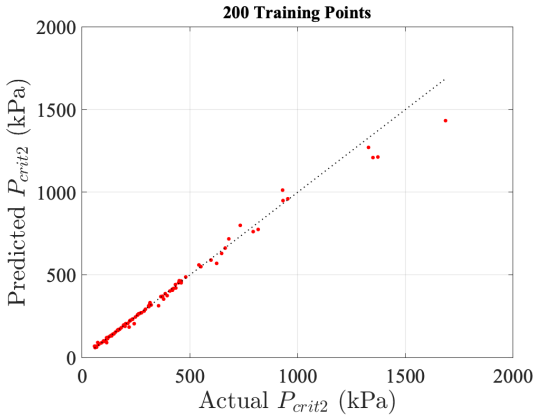


(b) 200 Training Points

Figure A.23: P_{crit1} Surrogate Predictions

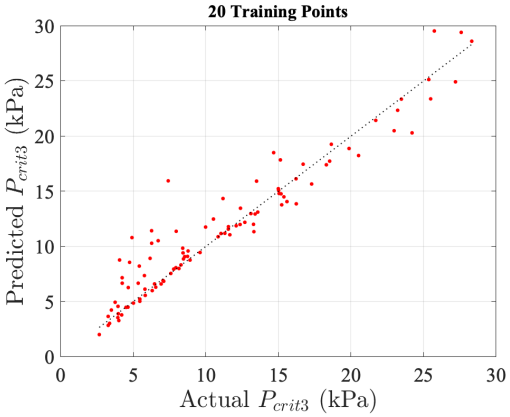


(a) 20 Training Points

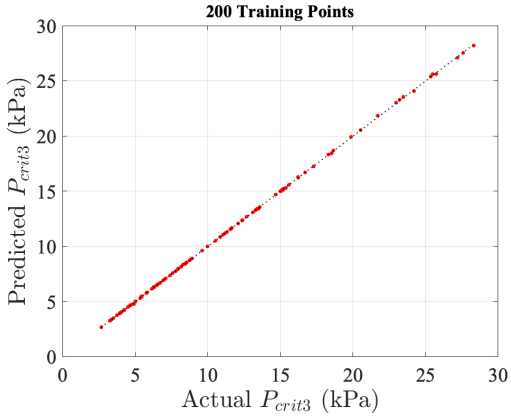


(b) 200 Training Points

Figure A.24: P_{crit2} Surrogate Predictions

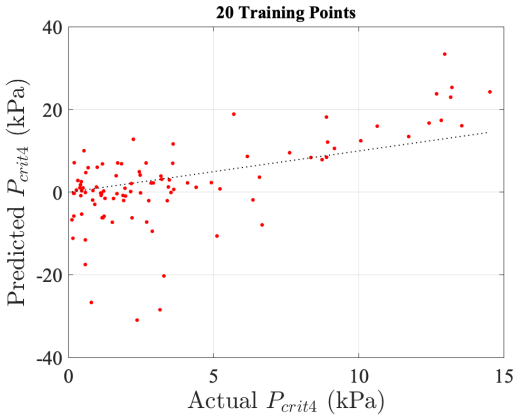


(a) 20 Training Points

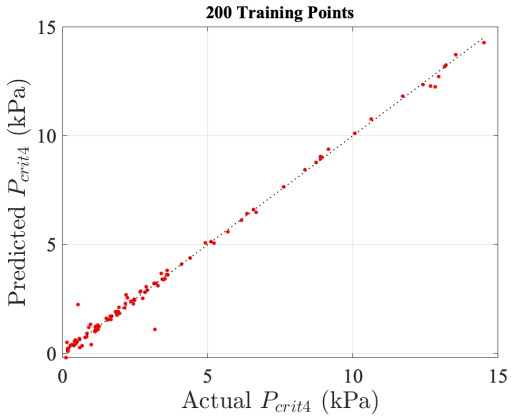


(b) 200 Training Points

Figure A.25: P_{crit3} Surrogate Predictions

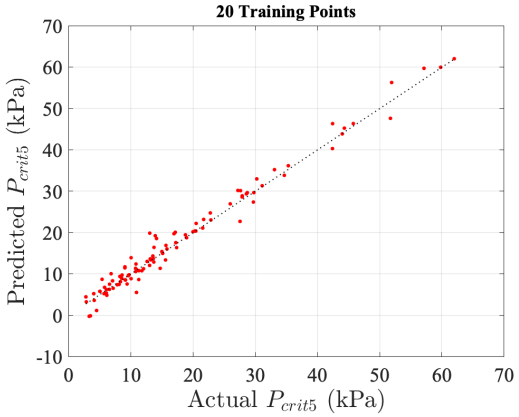


(a) 20 Training Points

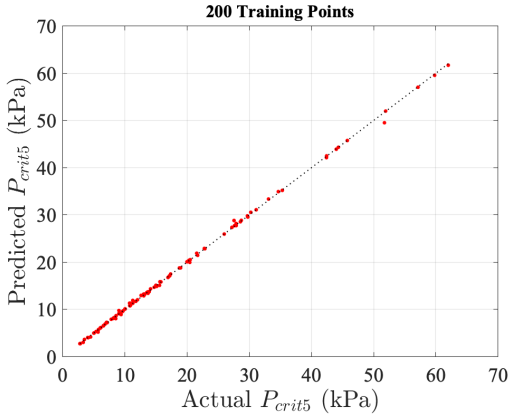


(b) 200 Training Points

Figure A.26: P_{crit4} Surrogate Predictions

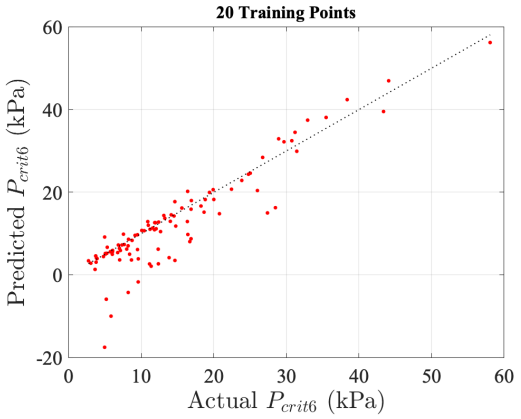


(a) 20 Training Points

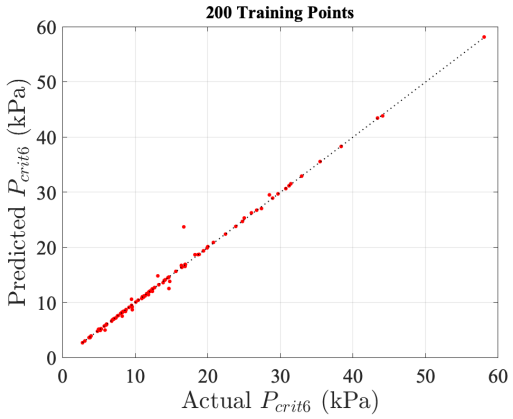


(b) 200 Training Points

Figure A.27: P_{crit5} Surrogate Predictions

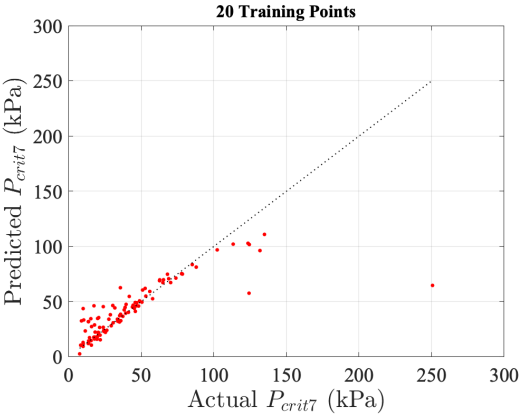


(a) 20 Training Points

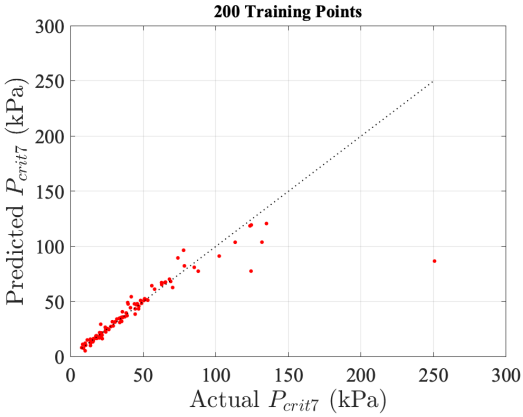


(b) 200 Training Points

Figure A.28: P_{crit6} Surrogate Predictions

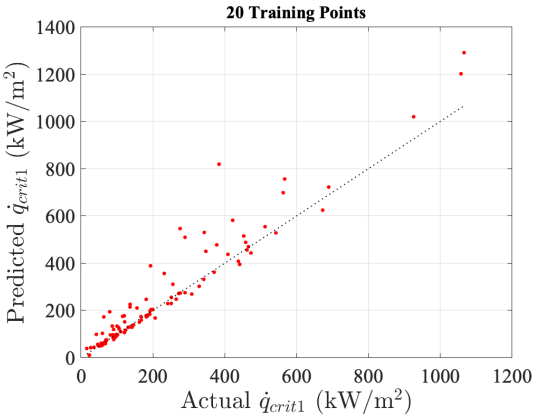


(a) 20 Training Points

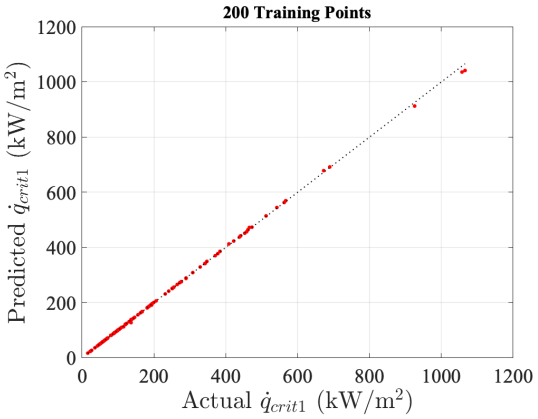


(b) 200 Training Points

Figure A.29: P_{crit7} Surrogate Predictions

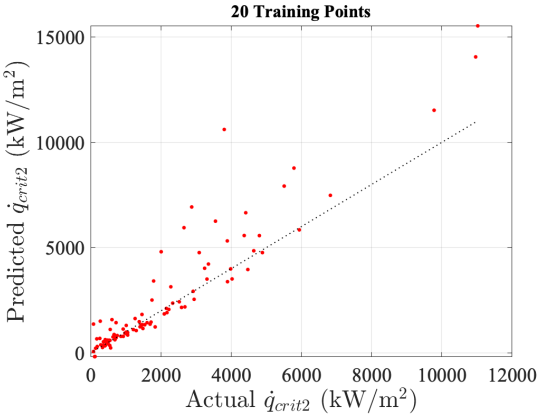


(a) 20 Training Points

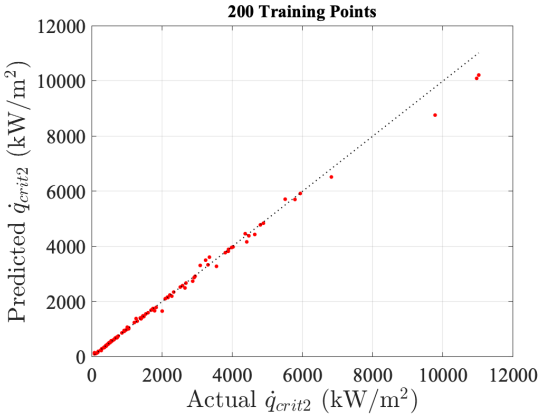


(b) 200 Training Points

Figure A.30: $\dot{q}_{crit1}$ Surrogate Predictions

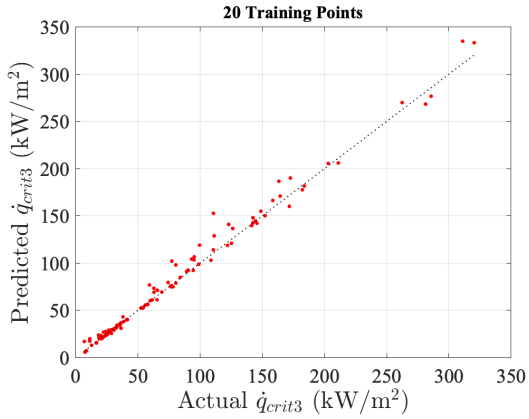


(a) 20 Training Points

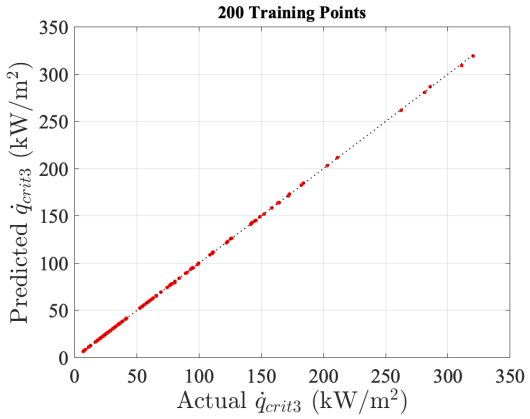


(b) 200 Training Points

Figure A.31: $\dot{q}_{crit2}$ Surrogate Predictions

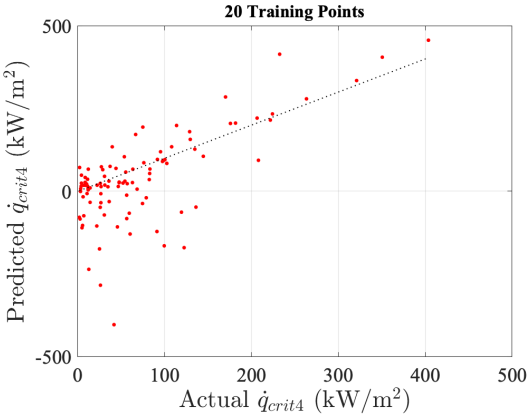


(a) 20 Training Points

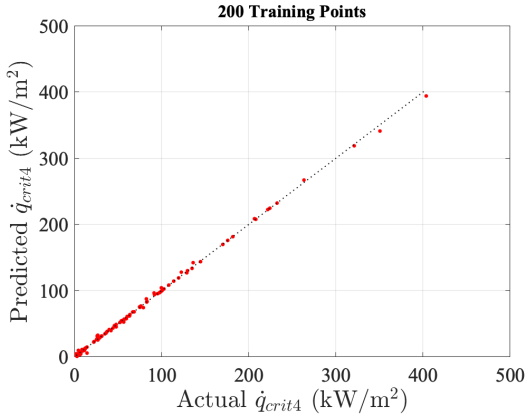


(b) 200 Training Points

Figure A.32: $\dot{q}_{crit3}$ Surrogate Predictions

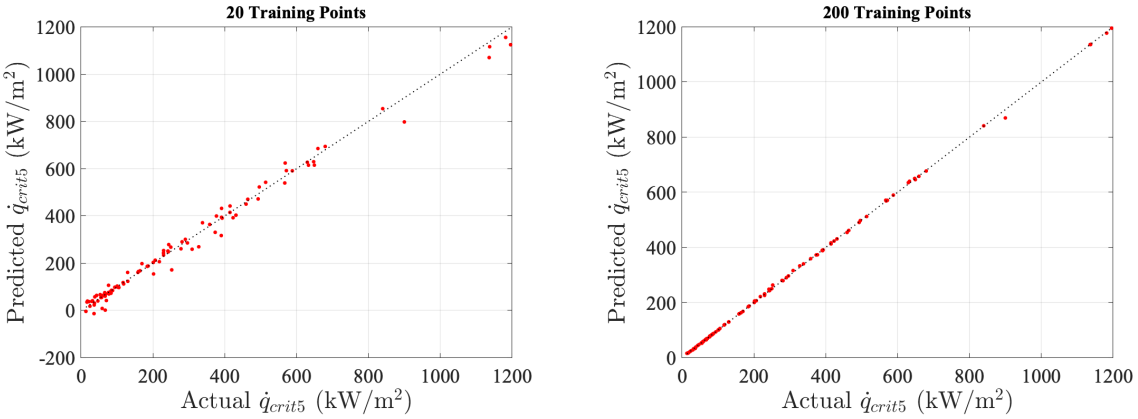


(a) 20 Training Points



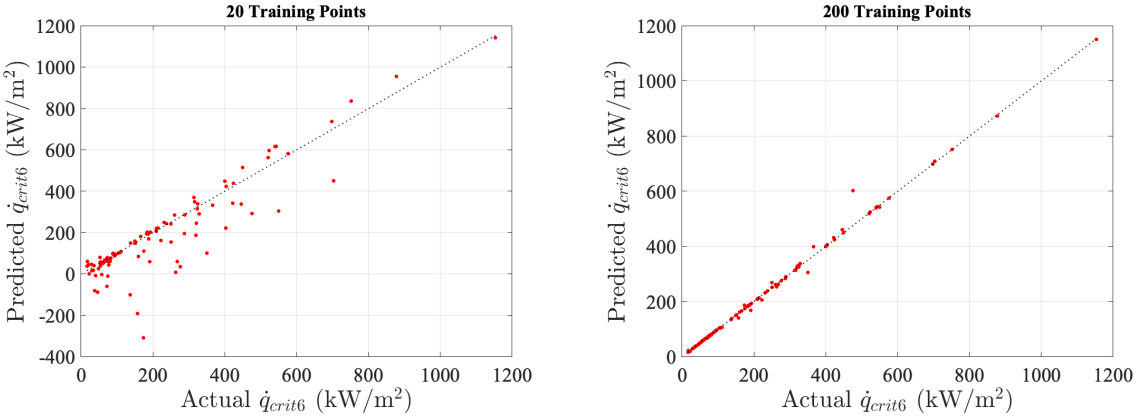
(b) 200 Training Points

Figure A.33: $\dot{q}_{crit4}$ Surrogate Predictions



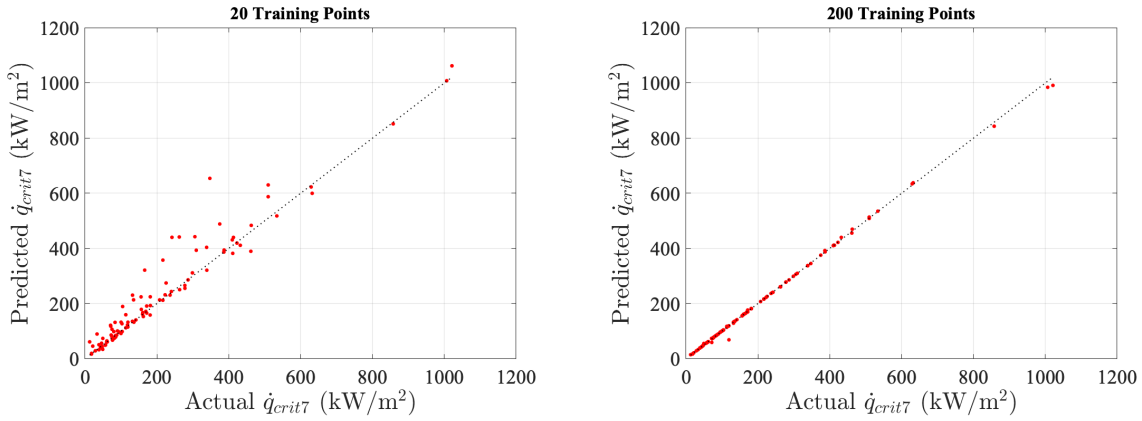
(a) 20 Training Points (b) 200 Training Points

Figure A.34: $\dot{q}_{crit5}$ Surrogate Predictions



(a) 20 Training Points (b) 200 Training Points

Figure A.35: $\dot{q}_{crit6}$ Surrogate Predictions



(a) 20 Training Points (b) 200 Training Points

Figure A.36: $\dot{q}_{crit7}$ Surrogate Predictions

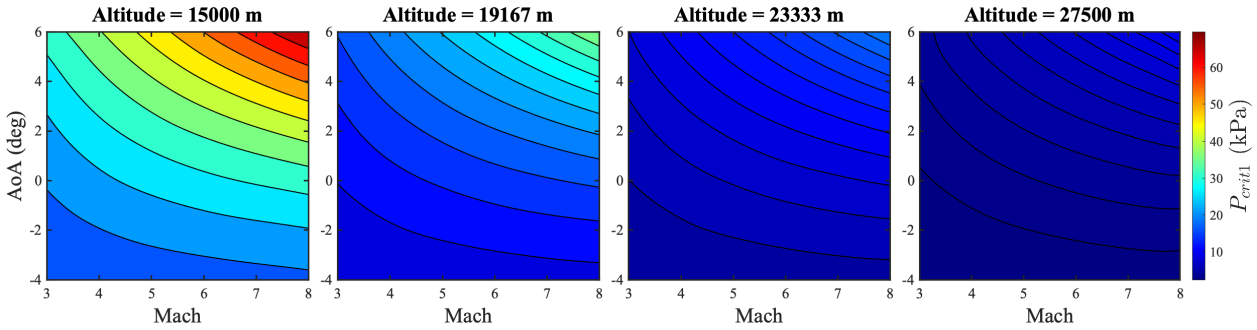


Figure A.37: P_{crit1} Surrogate Contours (900 Training Points)

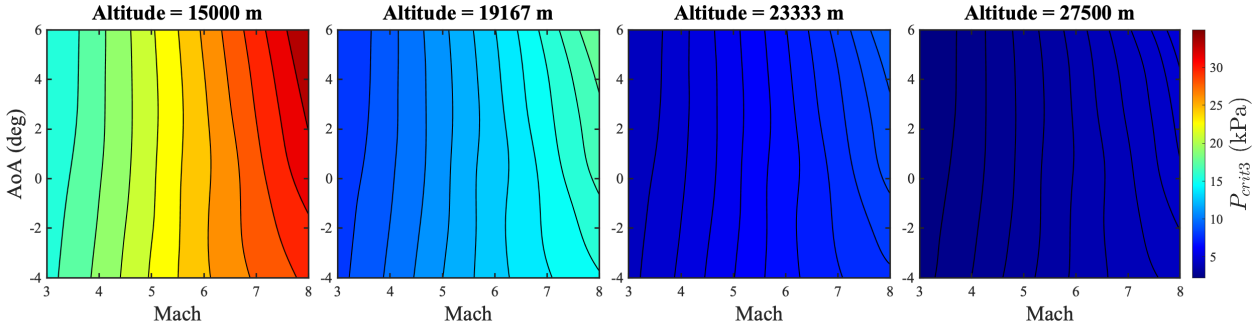


Figure A.39: P_{crit3} Surrogate Contours (900 Training Points)

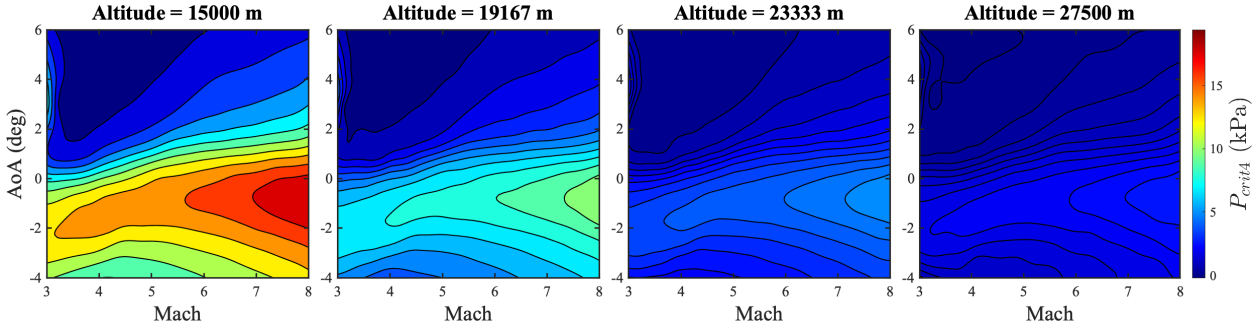


Figure A.40: P_{crit4} Surrogate Contours (900 Training Points)

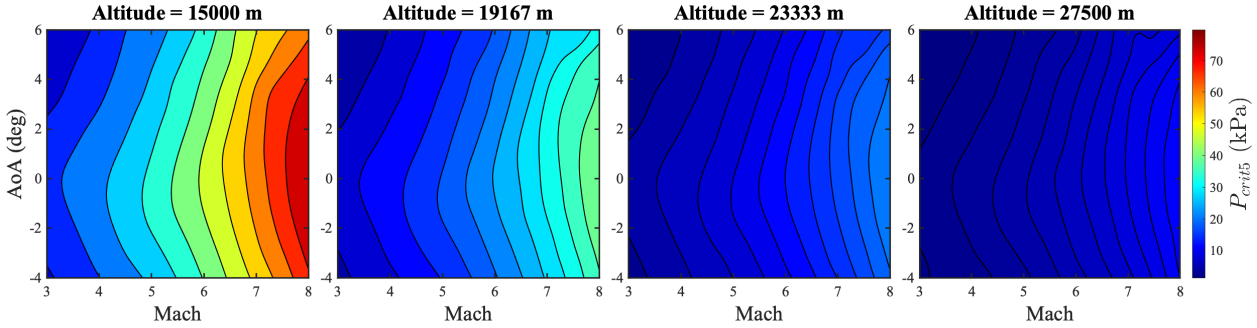


Figure A.41: P_{crit5} Surrogate Contours (900 Training Points)

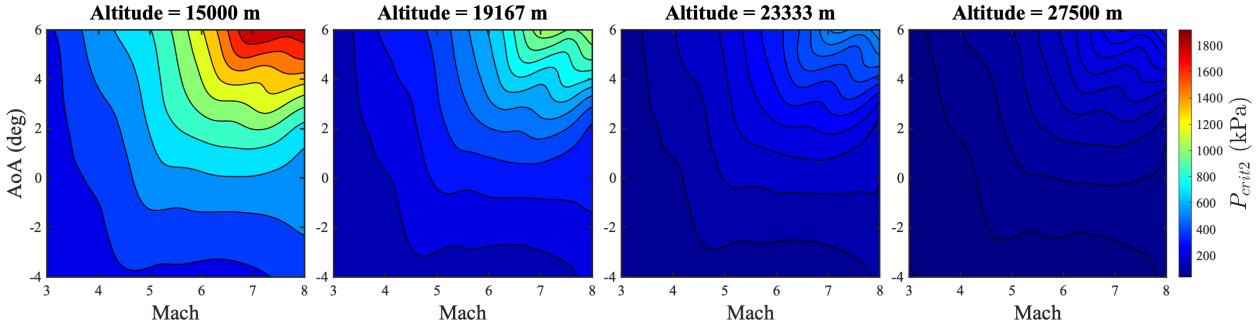


Figure A.38: P_{crit2} Surrogate Contours (900 Training Points)

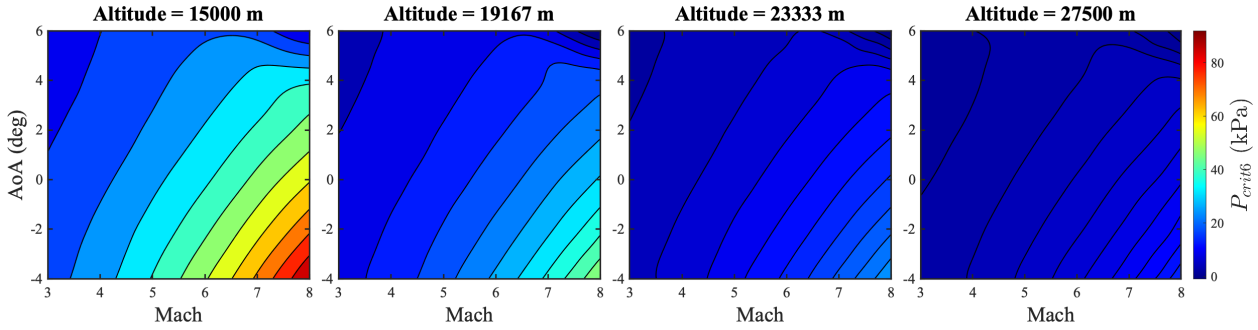


Figure A.42: P_{crit6} Surrogate Contours (900 Training Points)

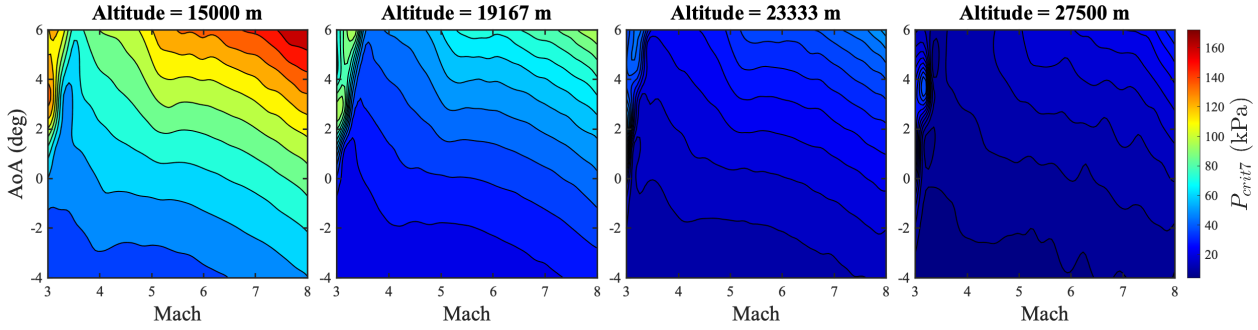


Figure A.43: P_{crit7} Surrogate Contours (900 Training Points)

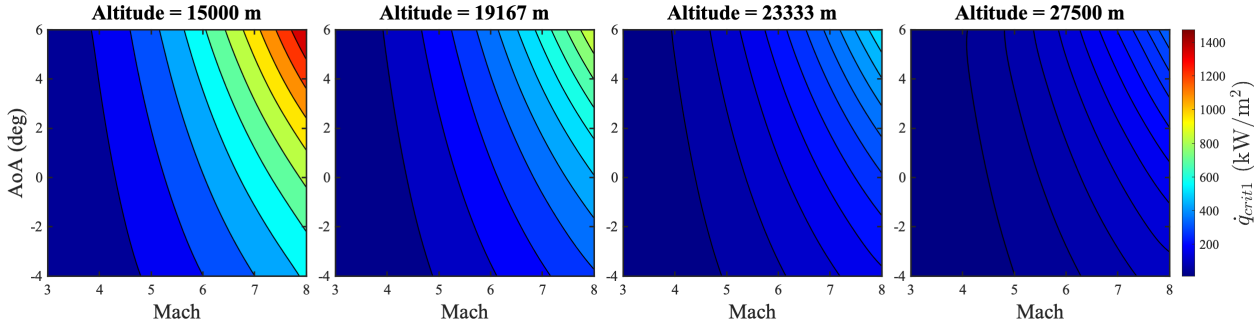


Figure A.44: $\dot{q}_{crit1}$ Surrogate Contours (900 Training Points)

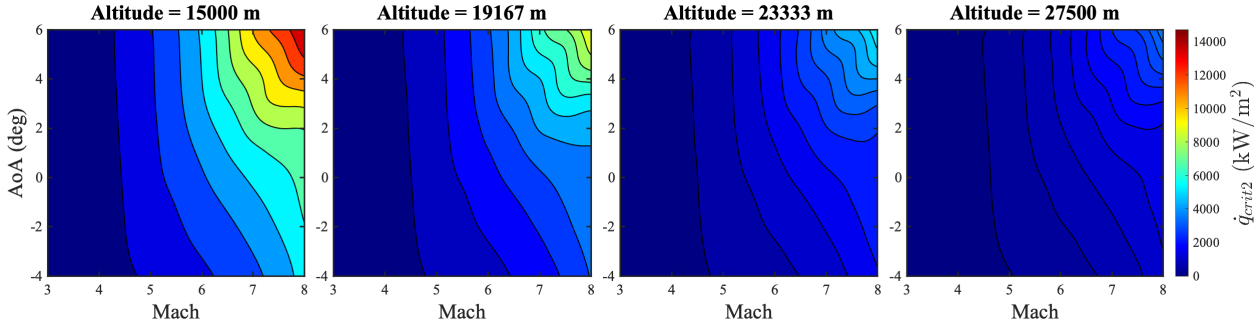


Figure A.45: $\dot{q}_{crit2}$ Surrogate Contours (900 Training Points)

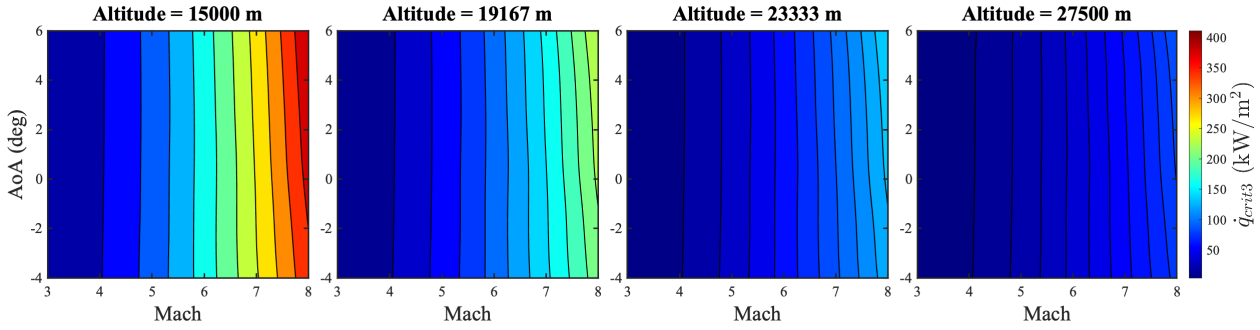


Figure A.46: $\dot{q}_{crit3}$ Surrogate Contours (900 Training Points)

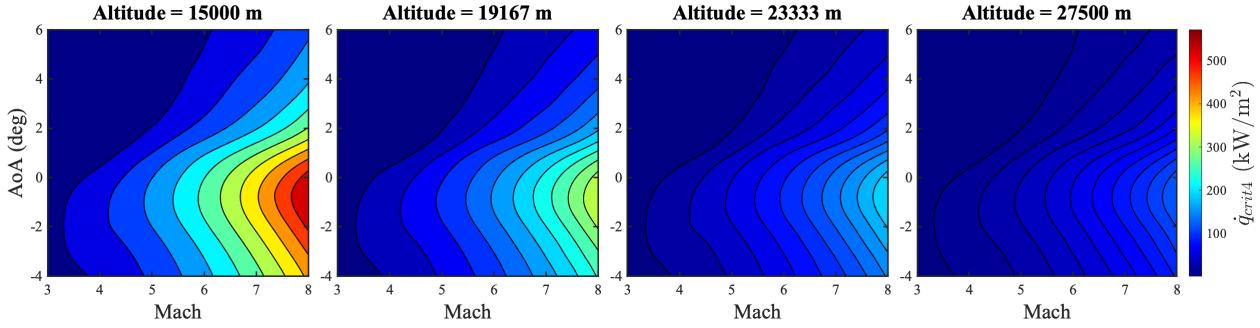


Figure A.47: $\dot{q}_{crit4}$ Surrogate Contours (900 Training Points)

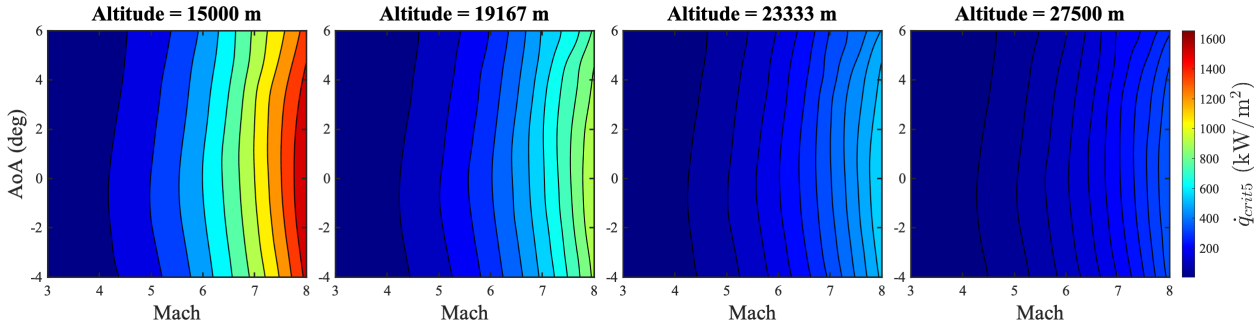


Figure A.48: $\dot{q}_{crit5}$ Surrogate Contours (900 Training Points)

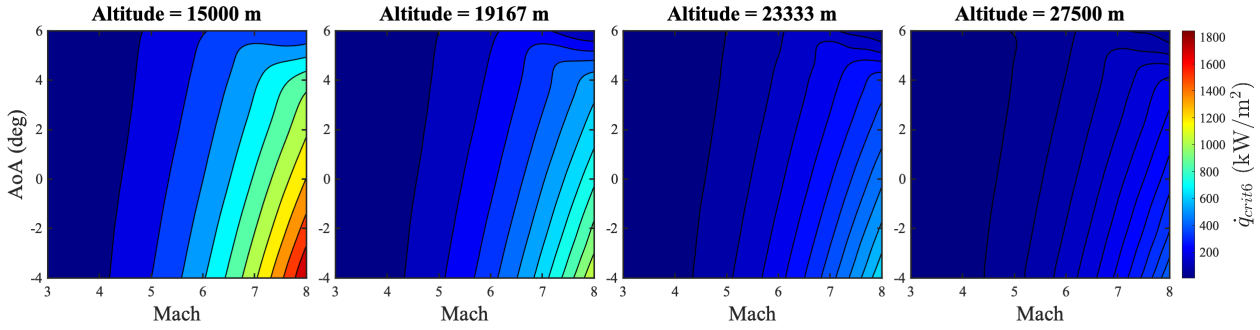


Figure A.49: $\dot{q}_{crit6}$ Surrogate Contours (900 Training Points)

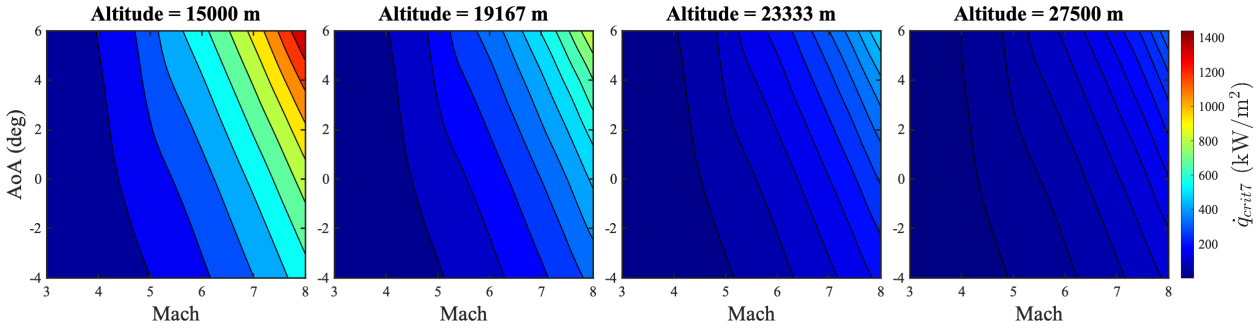
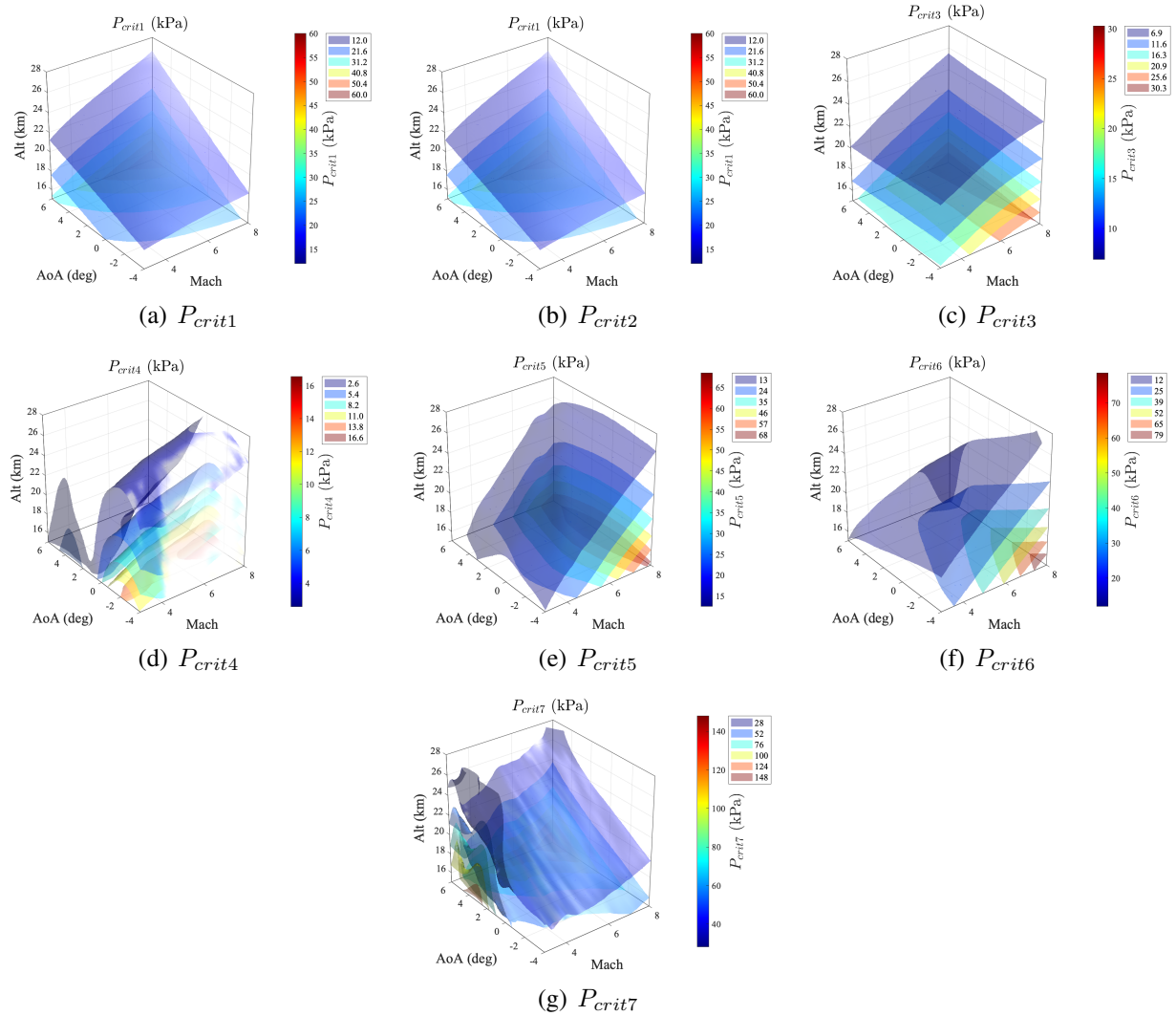


Figure A.50: $\dot{q}_{crit7}$ Surrogate Contours (900 Training Points)

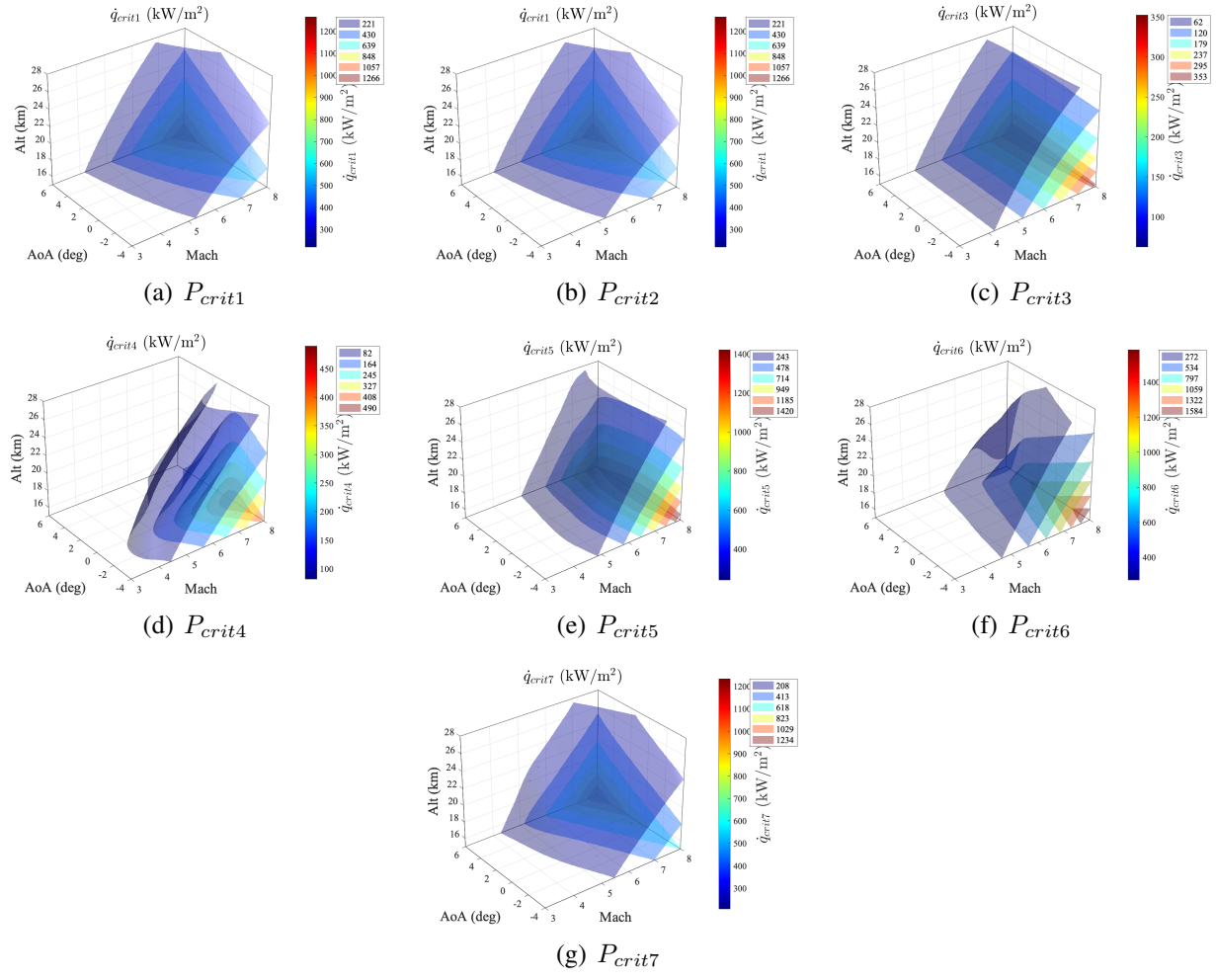


(a) P_{crit1} (b) P_{crit2} (c) P_{crit3}

(d) P_{crit4} (e) P_{crit5} (f) P_{crit6}

(g) P_{crit7}

Figure A.51: Critical Pressure Surrogate Isosurfaces (900 Training Points)



(a) P_{crit1} (b) P_{crit2} (c) P_{crit3}

(d) P_{crit4} (e) P_{crit5} (f) P_{crit6}

(g) P_{crit7}

Figure A.52: Critical Pressure Surrogate Isosurfaces (900 Training Points)