This electronic thesis or dissertation has been downloaded from Explore Bristol Research, http://research-information.bristol.ac.uk

Author:
Walsh, Jonathan A

Title:
A consideration of geometry in very-low Earth orbit satellites

General rights
Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

Take down policy
Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

•Your contact details
•Bibliographic details for the item, including a URL
•An outline nature of the complaint

Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.
This electronic thesis or dissertation has been downloaded from Explore Bristol Research, http://research-information.bristol.ac.uk

Author: Walsh, Jonathan A
Title: A Consideration of Geometry in Very-Low Earth Orbit Satellites

General rights
Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

Take down policy
Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

• Your contact details
• Bibliographic details for the item, including a URL
• An outline nature of the complaint

Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.
Abstract

There is increasing interest in operating Earth observation and telecommunication platforms at altitudes below 300km, in a region known as Very Low Earth Orbit (VLEO). VLEO provides several advantages over higher altitude orbits, such as reduced payload size and mass, as well as reductions in the payload power requirements, and improvements in the downlink data rate. Operating in VLEO, therefore, provides the opportunity to reduce the overall cost of a platform for a given mission.

Despite its advantages, operating in VLEO is not without its challenges. One such challenge is the higher levels of atmospheric drag a satellite experiences in these lower orbits. The drag can be compensated for using a dedicated propulsion system, however, this means that the operational life of the platforms will be limited by the amount of fuel it can carry. It is therefore desirable to find ways to maximize the life of the satellite as much as possible. This thesis presents work that examines aspects of a satellite’s geometry and identifies aeroshell profiles that will minimize the drag experienced by the platform in VLEO while maintaining its usability.

The approach taken for this work is to approximate the results of the Direct Simulation Monte Carlo (DSMC) simulations using a Radial Basis Function (RBF)-based surrogate model. The DSMC simulations are used to calculate the aerodynamic forces on the satellite’s body but are computationally expensive to run. By carefully selecting the sampling locations using an adaptive sampling approach and through effective interpolation of the data using RBF, the number of simulations required to explore the design domain can be limited. This means that a more complete examination of the parameter space of the aeroshell profiles could then be performed.

Using multi-objective optimization, a set of Pareto-optimal satellite geometries was identified. This set traded-off minimizing the drag against maximizing the internal volume of four aeroshell profiles: Blunted Wedge, Elliptical profile, Double-Conic nose and Rounded-Conic nose. When the aspect ratio of the body was fixed, it was seen that the Elliptical aeroshell profile generally performed the best of the four profiles tested, particularly for higher aspect ratio bodies. This is because the Elliptical profiles provided a good compromise between low drag and large volume. However, for a given internal volume it was seen that a Blunted Wedge profile was able to achieve a greater reduction in the drag experienced when compared to a similarly sized cuboid body. It was also shown that, in general, increasing the length of the satellite body to incorporate an aerodynamic profile is always beneficial.

Another factor to consider was the higher abundance of Atomic Oxygen (ATOX) in the VLEO environment. Since some payloads require access to space, they may be more susceptible to the corrosive nature of ATOX. The work presented approached this problem both analytically and with simulations as this helped to identify relationships and simplifications that could be applied in future work. It was also shown that the maximum ATOX flux inside a simple rectangular pit is experienced at the rim of the forward-facing panel and was approximately half the flux seen on the front of the satellite in the direct flow. It was shown that the angular distribution of particle through a point in space approximated a normal distribution in the orbital environment. This meant that the flux distribution on the forward-facing surface could be approximated by the cumulative frequency distribution of this normal distribution.
DEDICATION AND ACKNOWLEDGEMENTS

I would like to acknowledge the support of my academic supervisors, Lucy Berthoud and Chris Allen who have helped to guide the direction of my research.

I would also like to acknowledge my industrial colleagues at Thales Alenia Space UK for their help and support, in particular, Simon Chalkley, Andrew Bacon, Dawn Gilbert, & Piero-Francesco Siciliano.

This thesis is dedicated to my parents Anthony & Nicola Walsh and to my brother Daniel Walsh, who encouraged and supported me through this long journey.
This work was supported by an Engineering and Physical Sciences Research Council iCASE grant number 15220191 in partnership with Thales Alenia Space UK.

The views expressed in this thesis do not represent those of Thales Alenia Space UK.
I declare that the work in this thesis was carried out in accordance with the requirements of the University’s Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate’s own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the thesis are those of the author.

SIGNED: ........................................... DATE: 07/06/2022
# Table of Contents

<table>
<thead>
<tr>
<th>List of Tables</th>
<th>xii</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Figures</td>
<td>xv</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>xxiii</td>
</tr>
<tr>
<td>Nomenclature</td>
<td>xxvii</td>
</tr>
</tbody>
</table>

## 1 Introduction

1.1 Scope 1

1.2 Background 2

1.2.1 Definitions of Very Low Earth Orbit 2

1.2.2 Why operate in Very Low Earth Orbit (VLEO)? 4

1.2.3 Previous and planned visits to VLEO 5

1.2.4 Drag, Thrust and Fuel 7

1.2.5 Atomic Oxygen 8

1.3 Research Question and Derived Objectives 9

1.4 Structure of the Thesis 9

1.5 Publications 10

## 2 Initial exploration of satellite aerodynamics

2.1 Introduction 11

2.2 Method 12

2.2.1 Workflow 12

2.2.2 Central Body Geometry 13

2.2.3 Key Assumptions 14

2.3 The Atmosphere in VLEO 15

2.3.1 Overview 15

2.3.2 Models of the Atmosphere 15

2.3.3 Variations in the Neutral Atmosphere 16

2.3.4 Reduced Atmospheric Model 20

2.4 Modelling Aerodynamic Forces in VLEO 23

2.4.1 Overview 23

2.4.2 Molecular Flow 23

2.4.3 Methods for Determining Forces in a Molecular Flow 24
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.4.4 Gas-Surface Interaction Model</td>
<td>27</td>
</tr>
<tr>
<td>2.5 Result &amp; Discussion</td>
<td>30</td>
</tr>
<tr>
<td>2.5.1 Drag and Fuel</td>
<td>30</td>
</tr>
<tr>
<td>2.5.2 Sources of Drag</td>
<td>31</td>
</tr>
<tr>
<td>2.5.3 Results &amp; Discussion: Software Performance Comparison</td>
<td>35</td>
</tr>
<tr>
<td>2.6 Summary</td>
<td>37</td>
</tr>
<tr>
<td>3 Generation and Validation of a Surrogate Model for Rarefied Gas Modelling</td>
<td>39</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>39</td>
</tr>
<tr>
<td>3.2 Background</td>
<td>40</td>
</tr>
<tr>
<td>3.3 Components of a Surrogate Model</td>
<td>43</td>
</tr>
<tr>
<td>3.3.1 Interpolation</td>
<td>44</td>
</tr>
<tr>
<td>3.3.2 Sampling Strategy</td>
<td>46</td>
</tr>
<tr>
<td>3.3.3 The Function</td>
<td>51</td>
</tr>
<tr>
<td>3.3.4 Measuring Uncertainty</td>
<td>53</td>
</tr>
<tr>
<td>3.4 Initial Observations and Method Refinement</td>
<td>56</td>
</tr>
<tr>
<td>3.4.1 Choosing a Basis Function $\phi$</td>
<td>56</td>
</tr>
<tr>
<td>3.4.2 Choosing a Support Radius</td>
<td>58</td>
</tr>
<tr>
<td>3.4.3 Order of the Polynomial</td>
<td>61</td>
</tr>
<tr>
<td>3.4.4 Comparison of Sampling Strategies</td>
<td>62</td>
</tr>
<tr>
<td>3.4.5 Sample Budget Easing to improve RMSE</td>
<td>66</td>
</tr>
<tr>
<td>3.5 DSMC and Final Evaluation</td>
<td>69</td>
</tr>
<tr>
<td>3.5.1 Introduction</td>
<td>69</td>
</tr>
<tr>
<td>3.5.2 Method</td>
<td>69</td>
</tr>
<tr>
<td>3.5.3 Results &amp; Discussion</td>
<td>69</td>
</tr>
<tr>
<td>3.5.4 Summary</td>
<td>71</td>
</tr>
<tr>
<td>3.6 Conclusion</td>
<td>71</td>
</tr>
<tr>
<td>4 The Optimization of Satellite's Aerodynamics in VLEO</td>
<td>73</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>73</td>
</tr>
<tr>
<td>4.2 General Methodology</td>
<td>74</td>
</tr>
<tr>
<td>4.2.1 Workflow</td>
<td>74</td>
</tr>
<tr>
<td>4.2.2 Properties of the Fluid</td>
<td>77</td>
</tr>
<tr>
<td>4.2.3 Satellite Geometry</td>
<td>78</td>
</tr>
<tr>
<td>4.2.4 Surrogate Model</td>
<td>84</td>
</tr>
<tr>
<td>4.2.5 Optimization Strategy</td>
<td>88</td>
</tr>
<tr>
<td>4.3 Drag Sensitivity</td>
<td>92</td>
</tr>
<tr>
<td>4.3.1 Overview</td>
<td>92</td>
</tr>
<tr>
<td>4.3.2 Results &amp; Discussion: Blunted Wedge – Nose Profile</td>
<td>92</td>
</tr>
<tr>
<td>4.3.3 Results &amp; Discussion: Blunted Wedge – Tail Profile</td>
<td>96</td>
</tr>
<tr>
<td>4.3.4 Results &amp; Discussion: Elliptical Profile</td>
<td>102</td>
</tr>
<tr>
<td>4.3.5 Results &amp; Discussion: Double-Conic &amp; Rounded-Conic Nose</td>
<td>106</td>
</tr>
<tr>
<td>4.4 Fixed Body Dimensions</td>
<td>112</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS

4.4.1 Introduction ............................................. 112  
4.4.2 Results & Discussion ................................. 112  
4.4.3 Summary ................................................. 126  

4.5 Fixed Volume .............................................. 127  
4.5.1 Introduction ............................................. 127  
4.5.2 Fixed volume with fixed body length .................. 127  
4.5.3 Fixed body width ..................................... 131  
4.5.4 Summary ................................................ 136  

4.6 Future Work ................................................ 136  

4.7 Conclusion ................................................ 137  

5 Internal Environment in VLEO ........................... 139  
5.1 Introduction ................................................. 139  
5.2 General Method ............................................ 140  
5.2.1 Workflow ............................................... 140  
5.2.2 The Pit Geometry ..................................... 141  
5.2.3 Atmospheric Conditions ............................. 142  
5.3 Directional probability distribution of Gasses .............. 143  
5.3.1 Introduction ............................................. 143  
5.3.2 The Angular Distribution of Gas Species ............ 143  
5.3.3 Flux Distribution on a Wall .......................... 149  
5.3.4 Summary ................................................. 154  
5.4 Atomic Oxygen in a simple pit .......................... 156  
5.4.1 Atomic Oxygen Hot spots ............................. 156  
5.4.2 Depth Penetration ..................................... 162  
5.5 Conclusion ................................................ 164  

6 Conclusion .................................................. 165  
6.1 The Challenges of VLEO ................................. 165  
6.2 Surrogate Model .......................................... 166  
6.3 Shape Optimization ..................................... 167  
6.4 Atomic Oxygen .......................................... 171  
6.5 Future Research Topics ................................. 174  
6.6 Closing Remarks ......................................... 175  

A Notes on the development of the Rarefied Air Modeling .... 177  
A.1 A review of the precision and accuracy of the DSMC software SPARTA .... 177  
A.1.1 Introduction ............................................. 177  
A.1.2 Method ................................................. 178  
A.1.3 Results & Discussion ................................. 178  
A.1.4 Conclusion ............................................. 179  

B Notes on the development of the Surrogate Model .......... 183  
B.1 Inverting the system matrix of a Surrogate Model ...... 183
TABLE OF CONTENTS

B.2 Comparison of Sampling Strategies .................................................. 185
  B.2.1 Introduction .................................................................................. 185
  B.2.2 Method ......................................................................................... 185
  B.2.3 Result & Discussion ..................................................................... 185
  B.2.4 Summary ..................................................................................... 190

B.3 Choosing a Support Radius ................................................................. 192
  B.3.1 Introduction .................................................................................. 192
  B.3.2 Method ......................................................................................... 192
  B.3.3 Results & Discussion .................................................................. 193
  B.3.4 Summary ..................................................................................... 195

B.4 Order of the Polynomial .................................................................. 196
  B.4.1 Introduction .................................................................................. 196
  B.4.2 Method ......................................................................................... 196
  B.4.3 Result & Discussion .................................................................. 196
  B.4.4 Summary ..................................................................................... 198

B.5 Sample Budget Easing to improve RMSE ....................................... 199
  B.5.1 Introduction .................................................................................. 199
  B.5.2 When to relax the budget and when to stop .............................. 199
  B.5.3 Predicting the progress ................................................................. 200
  B.5.4 Extrapolating the Uncertainty ...................................................... 201
  B.5.5 The Break conditions ................................................................. 204
  B.5.6 Summary ..................................................................................... 205

C Derivations for the distributions of particles in a gas .................. 207
  C.1 The Angular Distribution of Gas Species ..................................... 207
    C.1.1 Introduction ............................................................................... 207
    C.1.2 Distribution of Velocities ......................................................... 207
    C.1.3 Molecular Speed Ratio ............................................................. 208
    C.1.4 Polar Form ............................................................................... 208
    C.1.5 Integrating the function ............................................................ 209
    C.1.6 Angular Deviation .................................................................. 211
  C.2 Integrating the Wall distribution ................................................... 213
    C.2.1 Element Analysis ..................................................................... 214
    C.2.2 Verifying the Simplification ..................................................... 215
    C.2.3 Integrating the Simplification .................................................. 217
    C.2.4 For large $s_\infty$ .................................................................... 218

References .............................................................................................. 219
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Two-dimensional candidate satellite profiles</td>
</tr>
<tr>
<td>2.2</td>
<td>Daily and 81-day average of the F10.7 solar index for the solar maximum and minimum used in this thesis</td>
</tr>
<tr>
<td>2.3</td>
<td>Properties of the atmosphere at 200km on the 1st March 2000 at 00:00:00 UTC as given by the spherically weighted mean of Mass Spectrometer and Incoherent Scatter Radar model produced by the US Naval Research Laboratory (NRL) (NRLMSISE-00) model using sampling points at 1° intervals in Latitude and Longitude</td>
</tr>
<tr>
<td>2.4</td>
<td>Molecular Species data required to implement the Variable Hard Sphere (VHS) model and used to calculate $\lambda_0$</td>
</tr>
<tr>
<td>2.5</td>
<td>Variation of Mean Free Path ($\lambda_0$) and Knudsen Number ($K_n$) with altitude for satellite with a reference length of 0.5m, Number Density from the last Solar Maximum (March 2000)</td>
</tr>
<tr>
<td>2.6</td>
<td>A Qualitative comparison of Molecular Flow solvers</td>
</tr>
<tr>
<td>2.7</td>
<td>Comparison of Aerodynamic Forces for each profile at 200km as well as estimated fuel requirements. Ref. values: $v = 7,794m/s$, $\rho = 5.30521 \times 10^{-10}kg/m^3$, $A_{ref} = 0.25m^2$</td>
</tr>
<tr>
<td>2.8</td>
<td>Comparison of Drag Coefficients from simulations performed using DS2V, Sentman and Sparta 2D &amp; 3D with errors relative to DS2V results</td>
</tr>
<tr>
<td>3.1</td>
<td>Common examples of radial basis functions</td>
</tr>
<tr>
<td>3.2</td>
<td>Design space boundary sample allocation and source per number of parameters</td>
</tr>
<tr>
<td>3.3</td>
<td>Number of monomial components required for a polynomial of order 0 to 5 for a design space with 1 to 5 variables</td>
</tr>
<tr>
<td>3.4</td>
<td>A comparison of sampling strategies with maximum sampling points of 100</td>
</tr>
<tr>
<td>3.5</td>
<td>$R^2$ Goodness-of-fit of a number curves to the RMSE values in log$_{10}$ space for each test case</td>
</tr>
<tr>
<td>3.6</td>
<td>Comparison of the model uncertainty after an initial sample budget guess of 100 and after applying the easing process for a required Root Mean Squared Error (RMSE) value of less than $1 \times 10^{-3}$</td>
</tr>
<tr>
<td>3.7</td>
<td>Comparison of the RMSE sample budget and run times for the Analytical and Direct Simulation Monte Carlo (DSMC) data sources for the Aero-1 and Aero-2 test cases</td>
</tr>
<tr>
<td>4.1</td>
<td>Properties of the atmosphere at 200km on the 1st March 2000 at 00:00:00 UTC as given by the spherically weighted mean of NRLMSISE-00 model using sampling points at 1° intervals in latitude and longitude</td>
</tr>
<tr>
<td>Table Reference</td>
<td>Title</td>
</tr>
<tr>
<td>-----------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>4.2</td>
<td>Typical dimensions for each satellite category used in the simulations</td>
</tr>
<tr>
<td>4.3</td>
<td>Parameters of the Blunted Wedge used in this chapter</td>
</tr>
<tr>
<td>4.4</td>
<td>Parameters of the Elliptical profile used in this chapter</td>
</tr>
<tr>
<td>4.5</td>
<td>Parameters of the Double-Conic nose used in this chapter</td>
</tr>
<tr>
<td>4.6</td>
<td>Parameters of the Rounded-Conic nose used in this chapter</td>
</tr>
<tr>
<td>4.7</td>
<td>Properties of the data set for the Blunted Wedge profile</td>
</tr>
<tr>
<td>4.8</td>
<td>Approximate number of samples needed by the adaptive sampling process used for the Blunted Wedge under similar atmospheric conditions but with varying number of parameters Note: assumes 16 cores</td>
</tr>
<tr>
<td>4.9</td>
<td>Geometric bounds on the design spaces and the profiles they apply to</td>
</tr>
<tr>
<td>4.10</td>
<td>Unconstrained optimized bodies for minimum $C_D$ for varying nose geometry and fixed (blunt) tail geometry using the Blunted Wedge profile</td>
</tr>
<tr>
<td>4.11</td>
<td>Unconstrained optimized bodies for minimum $C_D$ for varying tail geometry and fixed (blunt) nose using the Blunted Wedge profile</td>
</tr>
<tr>
<td>4.12</td>
<td>Approximate refill angles for gas species during the atmospheric condition described in Table 4.1</td>
</tr>
<tr>
<td>4.13</td>
<td>Unconstrained optimized bodies for minimum $C_D$ for varying nose geometry and fixed (blunt) tail geometry using the Elliptical profile</td>
</tr>
<tr>
<td>4.14</td>
<td>Unconstrained optimized bodies for minimum $C_D$ for varying tail geometry and fixed (blunt) nose geometry using the Elliptical profile</td>
</tr>
<tr>
<td>4.15</td>
<td>Unconstrained optimized bodies for minimum $C_D$ for varying nose geometry using Double-Conic profile</td>
</tr>
<tr>
<td>4.16</td>
<td>Unconstrained optimized bodies for minimum $C_D$ for varying nose geometry using Rounded-Conic nose</td>
</tr>
<tr>
<td>4.17</td>
<td>Example optimal Blunted Wedge profiles at varying points along the Pareto-optimal fronts for 1:2 profiles (left) and 1:6 profiles (right)</td>
</tr>
<tr>
<td>4.18</td>
<td>Example optimal Elliptical profiles at varying points along the Pareto-optimal fronts for 1:2 profiles (left) and 1:6 profiles (right)</td>
</tr>
<tr>
<td>4.19</td>
<td>Example optimal Double-Conic profiles at varying points along the Pareto-optimal fronts for 1:2 profiles (left) and 1:6 profiles (right)</td>
</tr>
<tr>
<td>4.20</td>
<td>Example optimal Rounded-Conic profiles at varying points along the Pareto-optimal fronts for 1:2 profiles (left) and 1:6 profiles (right)</td>
</tr>
<tr>
<td>4.21</td>
<td>Maximum % reduction in $C_D$ achievable with no net penalty assuming the fuel system volume scales with the drag value fixed at 1m</td>
</tr>
<tr>
<td>4.22</td>
<td>Comparison of the atmospheric drag of the Blunted Wedge and Elliptical profiles with the simple Cuboid profile given the desired volume and when the length of the profile is fixed at 1m</td>
</tr>
<tr>
<td>4.23</td>
<td>Comparison of the atmospheric drag of the Double-Conic and Rounded-Conic profiles with the simple Cuboid profile given the desired volume and when the length of the profile is fixed at 1m</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
</tr>
<tr>
<td>4.24</td>
<td>Comparison of the atmospheric drag of the Blunted Wedge and Elliptical profiles with the simple Cuboid profile given the desired volume and when the width of the profile is fixed at 0.5m</td>
</tr>
<tr>
<td>4.25</td>
<td>Comparison of the atmospheric drag of the Double-Conic and Rounded-Conic profiles with the simple Cuboid profile given the desired volume and when the width of the profile is fixed at 0.5m</td>
</tr>
<tr>
<td>5.1</td>
<td>Properties of the atmosphere at 200km on the 1st March 2000 at 00:00:00 UTC as given by the spherically weighted mean of NRLMSISE-00 model</td>
</tr>
<tr>
<td>5.2</td>
<td>Mean free speed ratio for the gas species to be simulated in subsequent DSMC simulations and their associated angular deviations $\sigma_{AD}$ and molecular Mach angles $\mu$</td>
</tr>
<tr>
<td>5.3</td>
<td>Comparison of maximum and ambient fluencies for each gas species on the Forward-facing panel (FFP) of a simple rectangular pit at 200km altitude Table 5.1</td>
</tr>
<tr>
<td>5.4</td>
<td>Key Atomic Oxygen (ATOX) flux values for a simple rectangular pit with openings of 0.1m, 0.3m and 0.5m wide and $\alpha_T$ of 0, 0.5 and 1. For reference $N_\infty = 4.732 \times 10^{19} \ m^{-2} \ s^{-1}$</td>
</tr>
<tr>
<td>6.1</td>
<td>Approximate number of samples needed by the adaptive sampling process used for the Blunted Wedge with varying number of parameters (Note: assumes 16 cores)</td>
</tr>
<tr>
<td>B.1</td>
<td>A comparison of sampling strategies with maximum sampling points of 100</td>
</tr>
<tr>
<td>B.2</td>
<td>Number of Monomial components required for a polynomial of order 0 to 5 for a design space with 1 to 5 variables</td>
</tr>
<tr>
<td>B.3</td>
<td>$R^2$ Goodness-of-fit of a number curves to the RMSE values in log_{10} space for each test case</td>
</tr>
<tr>
<td>B.4</td>
<td>Comparison of the model uncertainty after an initial sample budget guess of 100 and after applying the easing process for a required RMSE value of less than $1 \times 10^{-3}$</td>
</tr>
<tr>
<td>C.1</td>
<td>Mean speed ratio for the gas species and their associated angular deviations</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Comparison of the altitude ranges of LEO, VLEO, ELEO and SLEO as well as notable satellites that operate in or across these regions</td>
</tr>
<tr>
<td>1.2</td>
<td>Reasons for operating in VLEO and the primary drawbacks</td>
</tr>
<tr>
<td>1.3</td>
<td>GOCE and SLATS developed by ESA and JAXA respectively, both operate for prolonged periods within VLEO</td>
</tr>
<tr>
<td>2.1</td>
<td>Workflow for Chapter 2</td>
</tr>
<tr>
<td>2.2</td>
<td>Example configuration of VLEO satellites, note the fixed solar arrays parallel to the direction of flight</td>
</tr>
<tr>
<td>2.3</td>
<td>Variation of Atmospheric Density with Altitude, showing the range of densities during the last Solar Max (March 2000) and Solar Min (January 2008) according to the Atmospheric model NRLMSISE-00</td>
</tr>
<tr>
<td>2.4</td>
<td>F10.7 solar index from January 1986 to January 2021, showing the four recent solar cycles 22, 23, 24 &amp; 25. (Data obtained from the NASA Goddard Space Flight Center [44])</td>
</tr>
<tr>
<td>2.5</td>
<td>Density distribution during Solar Maximum and Solar Minimum</td>
</tr>
<tr>
<td>2.6</td>
<td>Relative density distribution during Solar Maximum and Solar Minimum</td>
</tr>
<tr>
<td>2.7</td>
<td>Variation of temperature with density across a range of altitudes in VLEO. Each cluster represents a single altitude during either high or low solar activity. The dashed line shows the mean temperature and density at each altitude during either high or low solar activity.</td>
</tr>
<tr>
<td>2.8</td>
<td>Variation in the composition of key gas species in the Thermosphere across a range of altitudes in VLEO. Each cluster represents a single altitude during either high or low solar activity. The dashed line shows the mean composition and density at each altitude during either high or low solar activity.</td>
</tr>
<tr>
<td>2.9</td>
<td>Atmosphere density from NRLMSISE-00 used for Drag modelling</td>
</tr>
<tr>
<td>2.10</td>
<td>Comparison of Drag Coefficients for the profiles at 200 km</td>
</tr>
<tr>
<td>2.11</td>
<td>Distribution of pressure and shear stress along the surfaces of Profiles S1 and S3 that contribute to the drag of the profile</td>
</tr>
<tr>
<td>2.12</td>
<td>Schematic view of the relative pressure (orange) and shear (green) for profile S1 and S3 (arrows not too scale)</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>2.13</td>
<td>Normalised density field around the profiles. Density field normalised to the free-stream density (red dotted line)</td>
</tr>
<tr>
<td>2.14</td>
<td>Comparison of Drag Coefficients from simulations performed using DS2V, Sentman and Sparta 2D &amp; 3D</td>
</tr>
<tr>
<td>3.1</td>
<td>Interactions of the components of a surrogate model</td>
</tr>
<tr>
<td>3.2</td>
<td>Example sampling strategies</td>
</tr>
<tr>
<td>3.3</td>
<td>Adding a new sample to the surrogate model using the adaptive criterion in Equation 3.19</td>
</tr>
<tr>
<td>3.4</td>
<td>Reference data for the four test cases</td>
</tr>
<tr>
<td>3.5</td>
<td>Geometry of the satellite body used in the aerodynamic test cases Aero-1 and Aero-2</td>
</tr>
<tr>
<td>3.6</td>
<td>Comparison of RMSE for each basis function with varying sampling budgets</td>
</tr>
<tr>
<td>3.7</td>
<td>Comparison of reciprocal condition number of the system matrix for each basis function with varying sampling budgets</td>
</tr>
<tr>
<td>3.8</td>
<td>Trade-off between reproducing the function and the conditioning of the systems matrix as the support radius changes</td>
</tr>
<tr>
<td></td>
<td>(Note: the size of the bubbles in the matrix indicate the amount of influence one point has over the other)</td>
</tr>
<tr>
<td>3.9</td>
<td>Comparison of RMSE for each sampling strategies with varying support radii and a fixed sample budget of 100</td>
</tr>
<tr>
<td>3.10</td>
<td>Reciprocal conditioning number for each sampling strategies with varying support radii and a fixed sample budget of 100</td>
</tr>
<tr>
<td>3.11</td>
<td>RMSE and run time to generate surrogate models with and without a polynomial</td>
</tr>
<tr>
<td>3.12</td>
<td>Comparison of RMSE for each sampling strategies with varying number of samples and fixed support radius $R_s = \sqrt{2}$</td>
</tr>
<tr>
<td>3.13</td>
<td>Generated surrogate model and absolute error in the model compared to the true function of the Droplet function for the 4 sampling strategies for $N = 100$</td>
</tr>
<tr>
<td>3.14</td>
<td>RMSE versus the total run time of the sampling strategies assuming DSMC runtime of 1 hour per sample</td>
</tr>
<tr>
<td>3.15</td>
<td>Comparison of the RMSE history for the four test cases up to a sample budget of 500 and assuming an initial 'guess' of 100</td>
</tr>
<tr>
<td>3.16</td>
<td>Surrogate Models for test case Aero-1 using the analytical rarefied gas equations developed by Sentman and the DSMC code SPARTA</td>
</tr>
<tr>
<td>3.17</td>
<td>Surrogate Models for test case Aero-2 using the analytical rarefied gas equations developed by Sentman and the DSMC code SPARTA</td>
</tr>
<tr>
<td>4.1</td>
<td>Overview of workflow for Chapter 4</td>
</tr>
<tr>
<td>4.2</td>
<td>Overview of functional flow for Chapter 4</td>
</tr>
<tr>
<td>4.3</td>
<td>Example candidate profiles to be examined in this section</td>
</tr>
<tr>
<td>4.4</td>
<td>Diagram showing the general form of the Blunted Wedge with key dimensions and direction of the velocity vector</td>
</tr>
<tr>
<td>4.5</td>
<td>Diagram showing the general form of the Elliptical profile with key dimensions and direction of the velocity vector</td>
</tr>
</tbody>
</table>
4.6 Diagram showing the general form of the Double-Conic nose with key dimensions and direction of the velocity vector .................................................. 82
4.7 Diagram showing the general form of the Rounded-Conic nose with key dimensions and direction of the velocity vector .................................. 83
4.8 Error history of the surrogate model data set used in this work for the Blunted Wedge profile ................................................................. 86
4.9 A slice through the surrogate model hyperspace, showing the variation of the $C_D$ with nose geometry with fixed tail geometry for the Blunted Wedge profile .................................................. 87
4.10 A slice through the surrogate model hyperspace, showing the variation of the $C_D$ with tail geometry with fixed nose geometry for the Blunted Wedge profile .................................................. 87
4.11 Optimum $C_D$ for varying nose length ($l_n$) with fixed (blunt) tail geometry using the Blunted Wedge profile .................................................. 94
4.12 Optimum $C_D$ for varying nose radius ($r_n$) with fixed (blunt) tail geometry using the Blunted Wedge profile .................................................. 94
4.13 Normalised density field and surface force vectors around the nose geometry of the $0.5m \times 1.0m$ Blunted Wedge profile. The density field is normalised to the free-stream density (red dotted line). Pressure vectors (red) and Force vectors (green) to scale and magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity. .................. 95
4.14 Optimum $C_D$ for varying tail radius ($r_t$) with fixed (blunt) nose geometry using the Blunted Wedge profile .................................................. 97
4.15 Optimum $C_D$ for varying tail length ($l_t$) with fixed (blunt) nose geometry using the Blunted Wedge profile .................................................. 97
4.16 Normalised density field and surface force vectors around the tail geometry of the $0.5m \times 1.0m$ Blunted Wedge profile. The density field is normalised to the free-stream density (red dotted line). Pressure vectors (red) and Force vectors (green) to scale and magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity. .................. 98
4.17 Variation of $C_D$ with half cone angle of the tail of the Blunted Wedge profile .................................................. 100
4.18 Optimum $C_D$ for varying nose length ($l_n$) with fixed (blunt) tail geometry using the Elliptical profile .................................................. 103
4.19 Optimum $C_D$ for varying tail length ($l_t$) with fixed (blunt) nose geometry using the Elliptical profile .................................................. 104
4.20 Normalised density field and surface force vectors around the nose geometry of the $0.5m \times 1.0m$ Elliptical profile. The density field is normalised to the free-stream density (red dotted line). Pressure vectors (red) and Force vectors (green) to scale and magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity. .................. 105
4.21 Sensitivity of the $C_D$ to the total length of the nose ($l_n$) .................................................. 107
4.22 Sensitivity of the $C_D$ to the front nose radius ($r_n$) .................................................. 108
4.23 Sensitivity of the $C_D$ to the front nose length of the Double-Conic profile ($l_f$) .................................................. 109
4.24 Normalised density field and surface force vectors around the nose geometry of the $0.5m \times 1.0m$ Double-Conic & Rounded-Conic nose. The density field is normalised to the free-stream density (red dotted line). Pressure vectors (red) and Force vectors (green) to scale and magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity. .................. 111
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.25</td>
<td>Graph showing the Pareto-optimal fronts for the Blunted Wedge profile when optimizing for minimum drag and maximum volume</td>
</tr>
<tr>
<td>4.26</td>
<td>Normalised density field and surface force vectors around the nose geometry of the 0.5m x 1.0m Blunted Wedge profile. The density field is normalised to the free-stream density (red dotted line). Pressure vectors (red) and force vectors (green) to scale and magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity.</td>
</tr>
<tr>
<td>4.27</td>
<td>Graph showing the Pareto-optimal fronts for the Elliptical profile when optimizing for minimum drag and maximum volume</td>
</tr>
<tr>
<td>4.28</td>
<td>Normalised density field and surface force vectors around the nose geometry of the 0.5m x 1.0m Elliptical profile. The density field is normalised to the free-stream density (red dotted line). Pressure vectors (red) and force vectors (green) to scale and magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity.</td>
</tr>
<tr>
<td>4.29</td>
<td>Graph showing the Pareto-optimal fronts for the Double-Conic profile when optimizing for minimum drag and maximum volume</td>
</tr>
<tr>
<td>4.30</td>
<td>Normalised density field and surface force vectors around the nose geometry of the 0.5m x 1.0m Double-Conic profile. The density field is normalised to the free-stream density (red dotted line). Pressure vectors (red) and force vectors (green) to scale and magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity.</td>
</tr>
<tr>
<td>4.31</td>
<td>Graph showing the Pareto-optimal fronts for the Rounded-Conic profile when optimizing for minimum drag and maximum volume</td>
</tr>
<tr>
<td>4.32</td>
<td>Normalised density field and surface force vectors around the nose geometry of the 0.5m x 1.0m Rounded-Conic profile. The density field is normalised to the free-stream density (red dotted line). Pressure vectors (red) and force vectors (green) to scale and magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity.</td>
</tr>
<tr>
<td>4.33</td>
<td>Variation of atmospheric drag with volume when the length of the body is fixed, comparing Cuboid profile (dashed line) to a Blunted Wedge (a), Elliptical profile (b), Double-Conic (c) and Rounded-Conic (d).</td>
</tr>
<tr>
<td>4.34</td>
<td>Improvement in atmospheric drag compared to a Cuboid profile when the length of the body is fixed.</td>
</tr>
<tr>
<td>4.35</td>
<td>Variation of atmospheric drag with volume when the width of the body is fixed, comparing Cuboid profile (dashed line) to a Blunted Wedge (a), Elliptical profile (b), Double-Conic (c) and Rounded-Conic (d).</td>
</tr>
<tr>
<td>4.36</td>
<td>Improvement in atmospheric drag compared to a Cuboid profile when the width of the body is fixed.</td>
</tr>
<tr>
<td>5.1</td>
<td>Overview of workflow in Chapter 5</td>
</tr>
<tr>
<td>5.2</td>
<td>Diagram showing the simulation set up for a simple rectangular pit, including key dimensions and the locations of the Forward-facing panel (FFP), Rearward-facing panel (RFP), Pit floor panel (PPF), Forward-facing rim (FFR) and Rearward-facing rim (RFR).</td>
</tr>
<tr>
<td>5.3</td>
<td>Diagram showing the velocity vectors and the projected angular probability distribution function $F(\theta)$.</td>
</tr>
<tr>
<td>5.4</td>
<td>Angular probability distribution of the molecular flux ($F(\theta)$) relative to the bulk fluid velocity ($v_\infty$).</td>
</tr>
</tbody>
</table>
5.5 Comparison of $\sigma_{AD}$ and $\mu$ with the numerical calculated standard deviation of $F(\theta)$.
5.6 Shielding effect criteria for concave corners from Koppenwaller et. al. [92].
5.7 Proportion of particles contained within a cone with a half angle of $3\sigma_{AD}$ and $2.5\mu$. 3SD cone from the numerical standard deviation included for comparison.
5.8 Angular probability distribution of the species in the atmosphere at an altitude of 200km under the condition provided in Table 5.1.
5.9 Angular number density distribution of the various species in the atmosphere at an altitude of 200km under the condition provided in Table 5.1.
5.10 A simple flat plate presented to a rarefied flow under steady-state conditions.
5.11 Diagram showing the inflow to a point $X$ on the surface of the FFP in a rectangular recessed pit.
5.12 Ratio of ambient flux in the pit $\tilde{N}_{pit}$ against simulation time.
5.13 Distribution of particle flux on the FFP of a recessed pit at 200km in altitude.
5.14 ATOX flux on the Forward-facing panel (FFP) (left column) and Rearward-facing panel (RFP) (right column) of a 0.1m wide pit for $\alpha_T = 0$, 0.5 and 1. $N_{\infty}$ and $\tilde{N}_{pit}$ included in the figure for reference.
5.15 ATOX flux on the Forward-facing panel (FFP) (left column) and Rearward-facing panel (RFP) (right column) of a 0.3m wide pit for $\alpha_T = 0$, 0.5 and 1. $N_{\infty}$ and $\tilde{N}_{pit}$ included in the figure for reference.
5.16 ATOX flux on the Forward-facing panel (FFP) (left column) and Rearward-facing panel (RFP) (right column) of a 0.5m wide pit for $\alpha_T = 0$, 0.5 and 1. $N_{\infty}$ and $\tilde{N}_{pit}$ included in the figure for reference.
5.17 Mean flux on the Pit floor panel (FFP) as a percentage of the external ambient flux with varying $\alpha_T$ and pit lengths of 0.1, 0.25 & 0.5m respectively.
6.1 Example candidate profiles to be examined in this section.
6.2 Graphs showing the Pareto-optimal fronts when optimizing for minimum drag and maximum volume.
6.3 Improvement in atmospheric drag using a Blunted Wedge profile compared to a Cuboid profile of similar volume when either the length or the width of the body is fixed.
6.4 Optimized aeroshell geometries for a 0.5m$^2$ body using a Blunted Wedge profile Original Cuboid drag = 13.80mN.
6.5 Diagram showing the velocity vectors and the projected angular probability distribution function $F(\theta)$.
6.6 Diagram showing the simulation set up for a simple rectangular pit, including key dimensions and the locations of the Forward-facing panel (FFP), Rearward-facing panel (RFP), Pit floor panel (FFP), Forward-facing rim (FFR) and Rearward-facing rim (RFR).
6.7 Distribution of Atomic Oxygen ($O_1$) flux on the FFP of a recessed pit at 200km in altitude.
A.1 Convergence of a set of similar Stochastic PArallel Rarefied-gas Time-accurate Analyzer (SPARTA) simulations on the S1 test body as the number of time steps is increased.
A.2 Convergence of a set of similar SPARTA simulations on the S1 test body as the number of simulated particles is increased.
LIST OF FIGURES
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATOX</td>
<td>Atomic Oxygen</td>
</tr>
<tr>
<td>CFD</td>
<td>Computational Fluid Dynamics</td>
</tr>
<tr>
<td>COTS</td>
<td>Commercial off-the-shelf</td>
</tr>
<tr>
<td>D-D</td>
<td>Dawn - Dusk Sun-synchronous Orbit</td>
</tr>
<tr>
<td>DSMC</td>
<td>Direct Simulation Monte Carlo</td>
</tr>
<tr>
<td>EGG</td>
<td>Electrostatic Gravity Gradiometer</td>
</tr>
<tr>
<td>EIF</td>
<td>Expected Improvement Function</td>
</tr>
<tr>
<td>ELEO</td>
<td>Extreme Low Earth Orbit</td>
</tr>
<tr>
<td>ESA</td>
<td>European Space Agency</td>
</tr>
<tr>
<td>FF</td>
<td>Fully Factorial</td>
</tr>
<tr>
<td>FFP</td>
<td>Forward-facing panel</td>
</tr>
<tr>
<td>FFR</td>
<td>Forward-facing rim</td>
</tr>
<tr>
<td>FOAM</td>
<td>Field Operation And Manipulation</td>
</tr>
<tr>
<td>GEO</td>
<td>Geostationary Orbit</td>
</tr>
<tr>
<td>GOCE</td>
<td>Gravity Field and Steady-State Ocean Circulation Explorer</td>
</tr>
<tr>
<td>GRACE</td>
<td>Gravity Recovery And Climate Experiment</td>
</tr>
<tr>
<td>GSIM</td>
<td>Gas-Surface Interaction Model</td>
</tr>
<tr>
<td>HS</td>
<td>Hard Sphere</td>
</tr>
<tr>
<td>ISS</td>
<td>International Space Station</td>
</tr>
<tr>
<td>JAXA</td>
<td>Japanese Aerospace Exploration Agency</td>
</tr>
<tr>
<td>LDEF</td>
<td>Long Duration Exposure Facility</td>
</tr>
<tr>
<td>LEO</td>
<td>Low Earth Orbit</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>LH</td>
<td>Latin Hypercube</td>
</tr>
<tr>
<td>LHS</td>
<td>Latin Hypercube Sampling</td>
</tr>
<tr>
<td>LOOCPV</td>
<td>Leave One Out Cross-Point Validation</td>
</tr>
<tr>
<td>LOX</td>
<td>Liquid Oxygen</td>
</tr>
<tr>
<td>LTAN</td>
<td>Local Time at Ascending Node</td>
</tr>
<tr>
<td>MEO</td>
<td>Medium Earth Orbit</td>
</tr>
<tr>
<td>MSISE</td>
<td>Mass Spectrometer and Incoherent Scatter Radar</td>
</tr>
<tr>
<td>N-M</td>
<td>Noon - Midnight Sun-synchronous Orbit</td>
</tr>
<tr>
<td>NASA</td>
<td>National Aeronautics and Space Administration</td>
</tr>
<tr>
<td>NRL</td>
<td>US Naval Research Laboratory</td>
</tr>
<tr>
<td>NRLMSISE-00</td>
<td>Mass Spectrometer and Incoherent Scatter Radar model produced by the <a href="#">NRL</a></td>
</tr>
<tr>
<td>OLHS</td>
<td>Optimized Latin Hypercube Sampling</td>
</tr>
<tr>
<td>PFP</td>
<td>Pit floor panel</td>
</tr>
<tr>
<td>PSO</td>
<td>Particle Swarm Optimization</td>
</tr>
<tr>
<td>RAAN</td>
<td>Right Accession of the Ascending Node</td>
</tr>
<tr>
<td>RAM</td>
<td>Rarefied Air Modelling</td>
</tr>
<tr>
<td>RBF</td>
<td>Radial Basis Function</td>
</tr>
<tr>
<td>RFP</td>
<td>Rearward-facing panel</td>
</tr>
<tr>
<td>RFR</td>
<td>Rearward-facing rim</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Squared Error</td>
</tr>
<tr>
<td>RSM</td>
<td>Polynomial Response Surface Model</td>
</tr>
<tr>
<td>RTM</td>
<td>Reduced Thermospheric Model</td>
</tr>
<tr>
<td>RTP</td>
<td>Ray-Tracing Panel</td>
</tr>
<tr>
<td>SLATS</td>
<td>Super Low Altitude Test Satellite</td>
</tr>
<tr>
<td>SLEO</td>
<td>Super Low Earth Orbit</td>
</tr>
<tr>
<td>SMG</td>
<td>Surrogate Model Generator</td>
</tr>
<tr>
<td>SOAR</td>
<td>Satellite for Orbital Aerodynamics Research</td>
</tr>
<tr>
<td>SPARTA</td>
<td>Stochastic PParallel Rarefied-gas Time-accurate Analyzer</td>
</tr>
<tr>
<td>SSO</td>
<td>Sun-synchronous Orbit</td>
</tr>
<tr>
<td>SST</td>
<td>Satellite-to-Satellite Tracking</td>
</tr>
<tr>
<td>TCM</td>
<td>Trajectory Correction Manoeuvre</td>
</tr>
<tr>
<td>TPMC</td>
<td>Test-Particle Monte Carlo</td>
</tr>
<tr>
<td>TSAT</td>
<td>Taylor Satellite</td>
</tr>
</tbody>
</table>
UCS  Union of Concerned Scientists
UKSA  UK Space Agency
VHS  Variable Hard Sphere
VLEO  Very Low Earth Orbit
VSS  Variable Soft Sphere
NOMENCLATURE

General

\( A_{\text{ref}} \)  Reference area  \\
\( C_D \)  Coefficient of Drag  \\
\( F_T \)  Thrust  \\
\( I_{sp} \)  Specific Impulse  \\
\( m_f \)  Fuel mass  \\
\( \rho \)  Density of the atmosphere  \\
\( T_0 \)  Mission Length  \\
\( v \)  Velocity  \\

Gaskinetic theory

\( \alpha_T \)  Energy accommodation coefficient  \\
\( d_{\text{ref}} \)  Reference particle diameter at \( T_{\text{ref}} \)  \\
\( E_i \)  Kinetic energy of incident molecule  \\
\( E_r \)  Kinetic energy of re-emitted molecule  \\
\( E_w \)  Kinetic energy of re-emitted molecule if re-emitted at the temperature of the wall  \\
\( K_n \)  Knudsen Number  \\
\( l_{\text{ref}} \)  Reference length  \\
\( \lambda \)  Mean free path  \\
\( m_s \)  Particle mass of gas species \( s \)  \\
\( n \)  Particle number density  \\
\( \omega \)  Viscosity-temperature index  \\
\( T \)  Gas Temperature  \\
\( T_{\text{ref}} \)  Reference Temperature
ABBREVIATIONS

**Surrogate Model (Chapter 3)**

$\beta$  
Coefficients of the basis matrix

$C$  
Adaptive criterion

$D(x,y)$  
Droplet function

$F$  
Matrix of the monomial results for each sample point $x_i$ such that $f_{i,k} = f_k(x_i)$

$f(x_*)$  
Vector of the monomial results for the $x_*$

$F(x,y)$  
Franke’s function

$\gamma$  
Coefficients of the polynomial matrix

$h(x)$  
Separation Function

$M$  
Number of monomial elements

$N$  
Number of samples of the design space

$n$  
Number of parameters of the design space

$p(x)$,  
Optional polynomial with $M$ monomial elements

$P_{\phi, \chi}(x_*)$  
The power function

$r(x_*)$  
Vector of the basis function for the $x_*$

$r$  
Scaled distance from the sample point

$R_s$  
Support Radius

$X$  
Set of $N$ Samples of the design space such that $X = x_1, x_2, ..., x_N$

$x_i$  
Set of $n$ parameters that define a Samples $i$

$x_*$  
Arbitrary evaluation point

$Y$  
Set of $N$ results of the sample set $X$ such that $Y = y_1, y_2, ..., y_N$

$y(x)$  
The original function to be approximated

$\hat{y}(x)$,  
Surrogate Model approximation of the original function

**Shape Optimization (Chapter 4)**

$\alpha$  
Linear weighting to obtain Pareto-optimal front

$C_V(G)$  
Constraining factor of the geometry $G$

$f_c$  
The cost function for optimisation

$f_{C_D}(G)$  
Coefficient of drag of the geometry $G$
ABBREVIATIONS

\( f_D(G) \) Atmospheric drag of the geometry \( G \)

\( f_V(G) \) The volume of the geometry \( G \)

\( G \) Satellite Geometry

\( l_b \) Length of profile body

\( l_f \) Front nose length of profile body

\( l_n \) Nose length of profile body

\( l_t \) Tail length of profile body

\( r_n \) Nose radius of profile body

\( r_t \) Tail radius of profile body

\( w_b \) Width of profile body

**Internal Flow (Chapter 5)**

\( c' \) Most probable thermal speed

\( d_p \) Depth of pit for internal flow simulations

\( \eta_f \) Angles from \( X \) to the **Forward-facing rim (FFR)** of the pit

\( \eta_r \) Angles from \( X \) to the **Rearward-facing rim (RFR)** of the pit

\( F(\theta) \) Angular probability distribution

\( f(\theta) \) Maxwell-Boltzmann distribution

\( l_p \) Length of pit for internal flow simulations

\( m \) Mass of particle

\( \mu \) Molecular Mach Angle

\( \dot{N}_\infty \) Ambient flux on the exterior surface parallel to flow

\( N_\infty \) Maximum flux on surfaces perpendicular to the flow

\( \dot{N}_{pit} \) Ambient flux on the exterior surface parallel to flow

\( N_S \) Flux of species \( S \)

\( n_S \) Number density of gas species \( S \)

\( \Omega_{X,S} \) Particle flux distribution along a wall in a rarefied flow

\( s \) Molecular speed ratio

\( s_\infty \) Molecular speed ratio of the bulk flow
ABBREVIATIONS

$s_r$ Radial speed ratio
$S$ Given gas species in mixture
$s_w$ Molecular speed ratio of the particle along the z-axis
$\sigma_{AD}$ Angular standard deviation
$T$ Temperature of the Gas
$\theta$ Direction of particle
$u$ Velocity in x-axis
$v$ Particle velocity
$v$ Velocity in y-axis
$v_\infty$ Free-stream Velocity
$v_r$ Radial Velocity Vector $[u, v]$
$\tilde{v}$ Internal (thermal) velocity
$w$ Velocity in z-axis
$X$ Point on the surface

Constants

$g_0$ Gravitational Constant $9.81 \text{ m/s}^2$
$k$ Boltzmann Constant $1.38064852 \times 10^{-23} \text{ m}^2\text{kg}\text{s}^{-2}\text{K}^{-1}$

Atmospheric gas species

$Ar$ Argon
$H$ Hydrogen
$He$ Helium
$N_1$ Atomic Nitrogen
$N_2$ Nitrogen
$O_1$ Atomic Oxygen
$O_2$ Oxygen
1.1 Scope

There is increasing interest in operating Earth observation platforms at a lower altitude, in a region referred to as Very Low Earth Orbit (VLEO). VLEO provides several advantages over higher altitude orbits, such as improved resolution or reduced payload size and mass, as well as reductions in the payload power requirements, and improvements in the data rate. It, therefore, provides the opportunity to reduce the overall cost of a platform for a given mission.

Despite its advantages, operating in VLEO is not without its challenges. One such challenge is combating the higher levels of atmospheric drag a satellite would experience in these lower orbits. While the drag can be compensated for using a dedicated propulsion system, the operational life of the platforms will still be limited. It is therefore desirable to find ways to maximize the life of the satellite as much as possible. A key focus of this thesis will therefore be to examine what parameters affect drag how it might be minimized.

Another factor to consider is the higher abundance of Atomic Oxygen (ATOX) in the VLEO environment. Since some payloads and other subsystems require access to space, they may be more susceptible to the corrosive nature of ATOX, which would ultimately limit their operational lives. This thesis will therefore also examine what impact this higher abundance of ATOX has on VLEO satellites and, in particular, it will look at locations that might be more at risk due to lower altitude operation.

This chapter will lay out the structure of the thesis. As background to this work, it will also seek to define the region known commonly as VLEO and identify its defining features. It will discuss the benefits and the risks of operating in VLEO as well as previous flown and planned missions to this region.
1.2 Background

1.2.1 Definitions of Very Low Earth Orbit

At the time of writing, a formal definition of Very Low Earth Orbit (VLEO) does not currently exist, however, it is broadly understood to describe the region of orbital altitudes closest to the Earth surface. As the orbital altitude decreases, the interaction between the satellite and the atmosphere increases, which may ultimately result in the deorbit of the satellite as the drag perturbations become dominant. Most sources agree that the key characteristic of the VLEO region is the increased impact the atmospheric drag has on the satellite which is often mission limiting or requires significant changes in traditional satellite system designs[1, 2, 3].

The increased impact of atmospheric drag has been commonly observed in satellites operating in the lower portion of Low Earth Orbit (LEO) which is commonly defined as those orbits with a mean altitude ranging from 200 km up to 2000 km. Depending on the definition, VLEO can be thought of as either a subset of LEO or as an overlapping region. Some authors define VLEO as those orbits with a mean altitude below 450 km[4, 5, 1]. Generally, at 450 km, the aerodynamic drag is strong enough to make a satellite's orbit decay in less than 5 years requiring regular boosts to maintain the orbit. An excellent example is the International Space Station (ISS) which must perform a boosting manoeuvre every 6 months to maintain its orbital altitude.

Other sources define VLEO as those orbits with a mean altitude below 300 km[6, 7, 8, 9]. Below 300km, the satellite experiences significantly more drag which generally means their orbits decay in less than a year. For these orbits to be of commercial value, with mission lengths greater than a year, the satellite must adopt aerodynamic architecture or utilize continuous drag compensation. An example of this is Gravity Field and Steady-State Ocean Circulation Explorer (GOCE) which adopted a long slender body with a low thrust ion thruster to overcome the drag at its operational altitude.

There are a few other definitions and names that define or overlap with the region broadly described by VLEO, however, these definitions are not in frequent use. Super Low Earth Orbit (SLEO) for instance has been used to define orbits of around 200km and lower[10, 11], placing it below LEO in some definitions. Extreme Low Earth Orbit (ELEO) on the other hand has been used to broadly describe the ‘uncharted ionosphere’ region between 120 and 300 km[9]. Figure 1.1 provides a visual comparison of the altitude ranges of LEO, VLEO, ELEO and SLEO as well as notable satellites that operate in or across these regions.

For this thesis, the upper boundary of VLEO will be defined as 300km, in line with the definition used by the industrial sponsor of this work, Thales Alenia Space.
1.2. BACKGROUND

Figure 1.1: Comparison of the altitude ranges of LEO, VLEO, ELEO, and SLEO as well as notable satellites that operate in or across these regions.
1.2.2 Why operate in VLEO?

As noted above, a key feature of the VLEO region is the significant influence of the atmosphere on the satellite which limits the operational life of traditional satellite architecture. It is therefore prudent to ask why this region should hold any interest for commercial or scientific operation and exploration? It has been shown in several studies that operating Earth observation satellites in a VLEO altitude can offer advantages over platforms in higher orbits.

One of the major benefits of operating in VLEO is the improvements to payload performance, particularly for earth observation. Earth observation satellites typically orbit between 500-800km in altitude, well above any of the definitions for VLEO previously discussed. By bringing the platform closer to the Earth, such as by orbiting within the VLEO region, it is possible to improve the ground resolution of the optical payload. Alternatively, for a set ground resolution, a much smaller optical instrument is required on a VLEO satellite comparable to a similar device on a platform in a higher orbit\cite{12, 13}. This has the advantage that the satellite platform can be much smaller to support the smaller instrument which in turn also reduces the platform cost\cite{1, 14, 15}.

The lower altitude also has benefits for radio-based payloads, such as radar and communications. For instance, the shorter distance between the satellite in VLEO and the ground reduces the latency of the signal and improves the link budget\cite{1}. As with the optical payloads, this also means a smaller antenna can be used with reduced transmission power, which saves on both the power and mass budget for the system. The lower altitude and closer horizon also allow for better reuse of the frequency bands.

VLEO provides other benefits for the satellite platform and operation. Crucially, the high drag environment of VLEO provides passive and quick disposal of the satellite at the end of life. This reduces the risk of collisions with space debris for other satellites operating in the region. The length of time that these derelict satellites remain in orbit will depend on the conditions of the atmosphere at the time, but conservatively this is still within 10 years rather than many decades. Since the satellite is lower in the atmosphere, it also benefits from the protection the slightly
thicker atmosphere provides against radiation. While the radiation environment in VLEO has not been fully explored, the added protection the atmosphere may afford coupled with the relatively short life expectancy of platforms could allow for the use of Commercial off-the-shelf (COTS) systems.[4]

1.2.3 Previous and planned visits to VLEO

Despite the challenges that the region can pose there have been a handful of excursions to altitudes in VLEO. Early examples include the US Keyhole and USSR Zenit reconnaissance satellites during the cold war. The perigees of these satellites dropped as low as 160 km to improve the ground resolution of the pictures they took. In this instance, the shorter operational life was also desirable since these satellites were limited by how much photographic film they could carry and this film had to be returned to Earth to be examined. These satellites often had highly elliptical orbits with apogees high above the VLEO to prevent them from returning to Earth too early. However, this also limited the regions of the Earth from which these satellites could return high-resolution imagery.

In some cases, it is desirable to operate fully within VLEO to improve coverage, ideally with an operational life greater than a year. To achieve this could require regular orbit-raising manoeuvres or adopting a low thrust drag compensation scheme to maintain altitude. An example of such a satellite is [GOCE](http://www.gocesatellite.com) developed by European Space Agency (ESA). The 1077kg Earth observation satellite was launched in 2009 to examine the Earth’s gravitational field.[16]. Operating at an altitude of 260 km, it employed an ion propulsion system to provide the necessary thrust to compensate for the drag it experienced. Under this regime, it was able to maintain orbit for 55 months using 40kg of Xenon fuel, before deorbiting within 2 weeks of fuel depletion. During the final year of operation, [GOCE](http://www.gocesatellite.com) descended further down into the atmosphere, holding at several altitudes to improve the quality of the data, before finally running out of fuel at an altitude of 229km. At each holding altitude, one of the QinetiQ T5 engines was once again employed to ensure drag-free operation and as would be expected the thrust required steadily increased as it descended. Upon the depletion of the Xenon fuel, [GOCE](http://www.gocesatellite.com) continued to operate nominally for just under a month, before contact was finally lost at 100km. During this final deorbit phase of the mission, [GOCE](http://www.gocesatellite.com) was able to collect valuable data about the operation of satellites in the VLEO environment. This included the performance of the attitude and orbital control system which continued to function nominally right up to the loss of contact. During this phase, the thermal impact of the atmosphere was also monitored with notable increases in temperature in components located towards the front of the satellite as the spacecraft descended.[17].

The [Japanese Aerospace Exploration Agency (JAXA)](http://www.jaxa.jp) has also sent a satellite to this region of orbit called the [Super Low Altitude Test Satellite (SLATS)](http://www.jaxa.jp) (also known as Tsubame) with the aim to study the atmospheric environment in VLEO and in particular the effects of atomic oxygen. Launched in late 2017, this 400kg satellite descended to its research orbit between 180-250km where it employed a hybrid chemical and electrical propulsion system to remain there for 90 days.[18][11]. At the time of writing, [JAXA](http://www.jaxa.jp) has formally concluded this mission and released
preliminary results. Interestingly, the preliminary results from their Atomic Oxygen experiments showed 44% lower flux than was expected based on the Mass Spectrometer and Incoherent Scatter Radar model produced by the US Naval Research Laboratory (NRL) (NRLMSISE-00). Additionally, the authors noted no significant degradation on the materials they tested, however, they do also note the short period over which the experiment was run before the mission concluded [19].

There have also been many successful CubeSat missions to investigate the VLEO region, such as Taylor Satellite (TSAT) launched on 18th April 2014. Released from an altitude of 325km, its primary purpose as it descended was to investigate the E and F region of the global Ionosphere in VLEO. It also sought to demonstrate the capabilities of this class of spacecraft for collecting valuable atmospheric data in this orbital region at a relatively low cost [9].

There are some upcoming missions scheduled to visit this region to help further the understanding of operating satellites in VLEO. Following the success of GOCE, ESA has commissioned studies into potential follow-on missions. One example presented by E. Canuto et al. describes a long-distance satellite formation operating in VLEO [21]. Gravity measurements using formation flying has been performed on past missions with notable examples including the Gravity Recovery And Climate Experiment (GRACE). The paper by E. Canuto et al. sought to highlight some of the issues of formation flying in VLEO with drag compensation and how the issues could be addressed. Other future programs focus on expanding our understanding of the VLEO environment to enable further utilization of the region. Satellite for Orbital Aerodynamics Research (SOAR) has been developed by the team behind DISCOVERER, a consortium of institutions from across Europe researching technologies for VLEO [22, 23, 24]. When launched, SOAR will test and validate the aerodynamic performance of materials in VLEO. This will provide valuable information about the gas-surface interactions of these materials which should improve the predictions of the aerodynamic forces on satellites operating in this region.

![Figure 1.3: GOCE and SLATS](image-url)
1.2.4 Drag, Thrust and Fuel

The previous section presented several satellites that operated or will operate, within the VLEO region. To maintain their orbital altitudes many of these platforms utilized a drag compensations system. Due to their high specific impulse ($I_{sp}$), electrical propulsion systems are the ideal choice for supporting long-duration drag compensation. However, they still require fuel which becomes the limiting factor on the operational life of the platform\cite{25} as was the case with both GOCE and SLATS. The fuel required is proportional to the thrust and by extension the drag that the satellite experiences \cite{15}, therefore it is desirable to minimize the drag as much as practicable to reduce the fuel mass fraction or to extend the life of the platform.

Historically, accurate determination of the aerodynamic forces on a spacecraft in LEO was rarely needed. This is because the forces experienced by the satellite as a result of the atmosphere were usually dominated by other orbital perturbations in LEO. Additionally, as is highlighted in the review papers by Llop et. al. \cite{1} and Prieto et. al. \cite{26}, there is a high level of uncertainty surrounding all aspects of the atmospheric drag, from determining the atmospheric conditions in VLEO to the interactions between the gas and the surfaces of the satellite body.

Early examinations of satellite drag formed part of a wider effort to deduce the properties of the upper atmosphere\cite{27}. These studies often assumed a drag coefficient of 2.2, based on the work and recommendations of Cook in 1965\cite{28}. This includes the work done by Jacchia in 1972, which lead to the development of one of the earliest neutral atmosphere models\cite{29, 30}. To this day, a coefficient of 2.2 is still a good first approximation for the drag in LEO.

A key source of uncertainty for determining the atmospheric drag in VLEO, or even in LEO, is the interaction between the sparse gases and the surfaces of the satellite. These interactions are normally defined by the energy accommodation coefficient, which determines how much energy is lost to the surface by a gas particle during an interaction. The energy accommodation coefficient has been shown to range in value from 0.86 to 1, based on empirical findings by Moe et. al.\cite{27} for satellites in the LEO. This indicates that the interaction between the gas and the surfaces is predominantly diffuse. Further work by Pilinski et. al.\cite{31} has demonstrated that the gas surface interactions in LEO are driven by the adsorption of atomic oxygen. They present a model for calculating the satellites accommodation coefficients based on observations of several satellites in LEO.

In subsequent work, Moe et. al.\cite{32} also highlighted the difficulty of using ground-based testing to determine the energy accommodation coefficient, suggesting that any such determinations should be made using observation of satellites in orbit. This is because, historically, laboratory experiments were unable to duplicate the orbital conditions experienced by the satellites. More recently, Roberts et al. \cite{22, 33} have been developing a new test facility that should be capable of mimicking the conditions observed in orbit. At the time of writing, the Rarefied Orbital Aerodynamics Research facility (ROAR) is still being commissioned, but once operational will be capable of simulating the free molecular flow and levels of atomic oxygen flux observed in VLEO.
CHAPTER 1. INTRODUCTION

1.2.5 Atomic Oxygen

The corrosive effects of Atomic Oxygen in LEO have been previously described in earlier work \[34, 35\]. The Atomic Oxygen chemically reacts with exposed surfaces forming oxides that can be volatile or mechanically fragile. This may lead to the degradation of the optical, thermal and mechanical properties of the exposed materials. Steps are often taken to mitigate these effects through careful material selection or by applying coatings where necessary \[36\]. This is more critical in VLEO since the abundance of Atomic Oxygen increases at lower altitudes.

Atomic Oxygen forms in the upper atmosphere through the photo-dissociation of molecular oxygen by ultraviolet light from the sun. The formation rate of Atomic Oxygen significantly exceeds its recombination or the formation of oxides so as a result, it is relatively abundant in VLEO accounting for between 30-60% of the gas composition. While the absolute abundance of the Atomic Oxygen may still be quite low in VLEO, as a direct result of the high velocities of the satellites, the incidence flux and collision energy are both high which lead to the corrosive effects observed \[37\].

Early examinations of the atomic oxygen environment relied on the return of material samples from LEO and were therefore limited to missions aboard one of the space shuttles \[35\]. These include the Long Duration Exposure Facility, a large cylindrical satellite deployed by the space shuttle to examine the long-term effects of the near-earth environment on materials. More recent studies have employed ground-based testing to test materials before they are used in space \[38, 39\]. Most of these studies focused on the LEO region, however, from atmospheric models such as NRLMSISE-00, the atomic oxygen flux in VLEO is expected to be significantly higher.

JAXA recently concluded a mission to examine the VLEO environment including an examination of the concentration and impact of Atomic Oxygen (see Figure 1.4). SLATS descended through the VLEO altitudes over the course of a few months before running out of fuel \[11\]. Interestingly, preliminary results from their Atomic Oxygen experiments showed 44% lower flux than was expected based on the NRLMSISE-00 atmospheric model. Additionally, the authors noted no significant degradation on the materials they tested, however, they do also note the short period over which the experiment was run before the mission concluded \[19\].

![Figure 1.4: Illustration of the deployed SLATS spacecraft and onboard Atomic Oxygen experiments][11]
1.3 Research Question and Derived Objectives

The research question for the work performed as part of this thesis can be formulated as:

*What are the specific implications on the aerodynamic design and system configuration for satellites operating in VLEO?*

From this question, the following objectives were derived:

- Identify the key challenges of operating a satellite in VLEO
- Examine the relationship between the satellite geometry and the atmospheric drag experienced
- Propose methods for drag reduction through the optimization of the aero design
- Investigate the influx of atomic oxygen to the interior of the satellite

1.4 Structure of the Thesis

The structure of the thesis is as follows:

This chapter, Chapter 1, presented an overview of the VLEO environment and why there is increased interest in operating in this region. Chapter 2 will explore in detail the properties of the atmosphere in VLEO and in particular the higher levels of atmospheric drag experienced in this region, a defining characteristic of VLEO. It will discuss how the atmosphere in VLEO can be modelled and the level of variability of the properties of the atmosphere in this region. It will also outline methods for calculating the aerodynamic forces in the rarefied gas of VLEO. Finally, this chapter will present initial investigations into the effects of the geometry on the drag, based on previous work performed within academia and industry and which will provide a foundation for later chapters.

Chapter 3 will outline work performed to develop a surrogate model specific for this project, which was necessary to fully explore the geometry design space. As will be explained, this is necessary due to the computational load of the aerodynamic tools. The chapter will provide an appraisal of the types of surrogate models available as well as a more detailed description of the Radial Basis Function (RBF) methodology used here. It will then show how this method was implemented and tested for this work and what measures were necessary to ensure a stable and accurate approximation.

In Chapter 4 the surrogate model is used to investigate the drag performance of a number of aeroshell profiles within a rarefied gas and will present optimized configurations under different criteria. Initially, the objective would be to minimize the drag the satellites experience to determine what impact these profiles have as part of a sensitivity analysis. A multi-objective optimization is then performed on each of the profiles to minimize the drag while maximizing the volume for a set of fixed external dimension constraints. Finally, this will be extended in Section 4.5 to optimize the aeroshell profiles for a given volume. Sections of this chapter have formed the basis of an Acta Astronautica paper published in January 2021[7].
The work presented in Chapter 5 will explore the internal environment of a satellite in Very Low Earth Orbit (VLEO) and in particular the ingress and propagation of Atomic Oxygen. This chapter was prompted by discussions on the refill rates of gases following the presentation of results in Chapter 4 and in particular with respect to the tapering of the tail sections of the satellite. The chapter will assess how atomic oxygen would be distributed throughout the cavity or pit that is located on the side of a satellite. This is achieved using both analytical methods and simulation-based using Direct Simulation Monte Carlo (DSMC).

1.5 Publications


2.1 Introduction

As discussed in Chapter 1, there is increasing interest in improving the capabilities and reducing the cost of Earth observation and telecommunication platforms by operating in orbits with much lower altitudes, in a region commonly referred to as VLEO. However, due to the higher atmospheric density of this region, satellites encounter significantly higher atmospheric drag which limits their operational life. It is therefore desirable to examine how this drag comes about, identify the driving factors and determine whether it is possible to minimize the drag a satellite experiences in this region of orbit.

This chapter will present an initial exploration of the atmospheric drag experienced by satellites in VLEO. Building upon previous work [12, 13, 2], this chapter will explore the elements that contribute to the drag. In particular, this will include a basic examination of how the geometry affects the drag using a set of candidate aeroshell profiles for a VLEO Microsatellite, setting the foundation for later chapters. To facilitate this, this chapter will also help to outline many of the definitions, concepts and topics that will be used throughout this thesis.
2.2 Method

2.2.1 Workflow

An overview of the workflow for Chapter 2 is provided in Figure 2.1. As can be seen on the right of Figure 2.1, to explore how the geometry affected the drag it was first necessary to identify what profiles are to be used in the analysis. Candidate aeroshell profiles will be described in Section 2.2.2 and are drawn from previous work performed within academia \cite{12, 13} and industry \cite{2}. As described in Section 2.3, a key characteristic of VLEO is the higher atmospheric density experienced by the satellites. For the analysis performed here, the reduced atmosphere model described in Section 2.3.4 will be used to provide the properties of the atmosphere in VLEO.

![Figure 2.1: Workflow for Chapter 2](image-url)
To calculate the aerodynamic forces on the body of the satellite it is necessary to simulate the gas in VLEO. Section 2.4 will examine the properties of the fluid in VLEO, as well as present methods that are available for determining the aerodynamic forces in this medium.

The aerodynamic forces experienced by these candidate profiles will be presented in Section 2.5.1. To facilitate the analysis, the satellite was assumed to be in a circular orbit of 200km, placing it roughly in the centre of the VLEO region. This section will also provide an estimate of the fuel requirements of these profiles for a range of specific impulses.

A further consideration for the aerodynamic forces is presented in Section 2.5.2, which will analyse the distribution of aerodynamic forces across the candidate profiles. This will show how each of the surfaces contributes to the overall forces that are experienced by the profile which in turn will help identify areas of further improvement.

In Sections 2.5.1 & 2.5.2, the DSMC software Stochastic PArallel Rarefied-gas Time-accurate Analyzer (SPARTA) was used to calculate the forces on the candidate profiles. Section 2.5.3 will compare the results from SPARTA with similar results collected using the DS2V (often considered the gold standard of DSMC) and panel methods.

### 2.2.2 Central Body Geometry

Previous work [12, 13, 2] has shown that satellites in the microsatellite category (10-100kg) are commercially feasible in VLEO, so, for the work performed here, a 100kg VLEO microsatellite will be assumed as the baseline. This is significantly smaller than either the ESA satellite GOCE or the Japanese SLATS satellite, which were small and minisatellites respectively, but presents a good starting point for the analysis. This mass budget includes the mass of the payload as well as the mass of the fuel required for drag compensation. To be consistent with previous work [12, 13, 2], the body of the satellite has exterior dimensions of 1.0m x 0.5 m x 0.5 m with a reference area of 0.25 $m^2$. (For 3D simulations the profile is assumed to be constant across the width of the satellite).

The profiles used here consist of a set of symmetric geometries with varying nose geometries as shown in Table 2.1. The key aim of investigating these was to determine if the drag could be reduced by using a leading tapered section in a free-molecular flow. They also provide simple body shapes to help compare software capabilities, as well as to facilitate the understanding of the underlying physics.

<table>
<thead>
<tr>
<th>Profile</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume fraction [-]</td>
<td>1.00</td>
<td>0.90</td>
<td>0.88</td>
<td>0.80</td>
</tr>
<tr>
<td>Cone Length [m]</td>
<td>0.0</td>
<td>0.2</td>
<td>0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>Cone Angle [°]</td>
<td>–</td>
<td>45.0</td>
<td>33.7</td>
<td>21.8</td>
</tr>
</tbody>
</table>
2.2.3 Key Assumptions

1. The Centre of Mass is at the centre of the bounding box of the body
The distribution of mass within the spacecraft is currently unknown, therefore the centre of mass will be taken to be the centre of the bounding box of the body (where the bounding box implies the smallest cuboid that can contain the test body of the satellite, equivalent to the centre of profile S1).

2. Solar panels are body mounted
The size of the solar panels is dependent on several factors including the power requirements for the electrical propulsion system. Solar panels can either be body-mounted or actuated to face the sun. Choosing the latter has two implications: firstly, the angle that the array is at will have an impact on the calculated drag. Ascertaining the full extent of the drag effect would require a geometry with variable angles of solar panel, which is currently beyond the scope of this investigation. Secondly, near the poles, the solar arrays would present their full area to the flow and thus would have a large drag penalty. Thus, the solar panels are assumed to be body mounted such that only the body shapes described above contribute to the drag. Figure 2.2 shows example configurations of previous VLEO satellites: GOCE (Figure 2.2a) and SLATS (Figure 2.2b) which both feature body-mounted solar panels parallel to the flow direction.

3. No holes in the aeroshell or externally mounted equipment
Some internally mounted equipment such as the payload and startrackers will require holes to the exterior to operate. For the purposes of the simulations carried out here these holes will not be considered.

Figure 2.2: Example configuration of VLEO satellites, note the fixed solar arrays parallel to the direction of flight
2.3 The Atmosphere in VLEO

2.3.1 Overview

As described in Section 1, VLEO defines the region of the near-Earth environment below 300km. The defining feature of this region (common among all definitions of VLEO) is the increased impact atmospheric drag has on the satellite which often limits the length of the mission or requires significant changes compared to traditional satellite system designs[1, 2, 3]. It is therefore important to understand the properties of the atmosphere in this region to make valid predictions about the drag that the satellite may experience.

2.3.2 Models of the Atmosphere

To obtain a realistic estimation of the drag a satellite may experience it is important to obtain an accurate representation of the atmospheric properties. To achieve this, there are several models available but the two most commonly used for modelling the drag on a satellite are the Jacchia-Bowman Thermospheric Density Model 2008 (JB-2008) and the NRLMSISE-00[41]. These belong to the Jacchia and MSISE families of models respectively which, through successive iterations and improvements, have accrued significant heritage within the space industry.

The Jacchia family of models are principally based on observations of decaying satellites and were one of the earliest models developed to describe the atmosphere[29]. Subsequent versions have taken account of ground-based observations and as well as solar irradiance which resulted in the development of the more recent JB-2008 model[42]. JB-2008 is considered to be a fairly accurate model for total density between periods of Solar Maxima, with an estimated uncertainty of between 10% and 15%[41]. However, the uncertainty in the predictions made by JB-2008 increases significantly during periods of solar maxima due to a lack of supporting data.

NRLMSISE-00 is an atmospheric model produced by the NRL and is the latest in the Mass Spectrometer and Incoherent Scatter Radar (MSISE) model family. As its name suggests, the MSISE family consists of an empirical model built on ground-based observations from incoherent scatter radar. More recent iterations, such as the NRLMSISE-00, now also draw on a large number of other data sources such as the drag experienced by rockets during launch and satellites in orbit and on re-entry[30]. The model is capable of describing the neutral temperature as well as the total density and molecular composition at altitude. Compared to JB-2008, the NRLMSISE-00 model is less accurate with an estimated uncertainty of 15% or higher[41].

For the work performed as part of this thesis, the atmospheric model NRLMSISE-00 was used. Despite the lower accuracy, it has been the industry standard for a long time and is the recommended model for simulating the drag on satellite according to the European Space Agency Standard ECSS-E-ST-10-04C[41].
2.3.3 Variations in the Neutral Atmosphere

The orbital altitudes of VLEO fall within the lower portion of the Earth’s Thermosphere (between 120km - 400km altitude). Figure 2.3 shows the variation of atmospheric density with altitude within VLEO and the lower portion of LEO during the Solar Maximum in March of 2000 and Solar Minimum in January of 2008 according to the atmospheric model NRLMSISE-00. Also shown are the operational altitude ranges of 4 satellites that have operated at or near the VLEO region. As demonstrated by Figure 2.3 there is significant variation in the properties of the thermosphere in VLEO. For a given altitude, the properties of the atmosphere can change due to influences of the day/night cycle as well as the Earth’s ground topology, however, the major influence on the properties of the atmosphere is the 11-year solar cycle[43].

Figure 2.3: Variation of Atmospheric Density with Altitude, showing the range of densities during the last Solar Max (March 2000) and Solar Min (January 2008) according to the Atmospheric model NRLMSISE-00
2.3.3.1 Solar Cycle

Figure 2.4 shows the F10.7 solar index between 1986 and 2021 along with four of the most recent solar cycles (the current cycle is Cycle 25 which began in December 2019). This data was obtained from the NASA Goddard Space Flight Center through their online tool for heliospheric data. The F10.7 solar index is one of a number of indices used to measure the activity of the sun (the other common one being the sunspot count which is used to define the Solar Cycles). The solar index is commonly used as an input to atmospheric models, including NRLMSISE-00, which is why it is of interest here.

To help facilitate the analysis in this thesis it was necessary to identify a set of dates that are representative of solar maximum and minimum. The most recent completed cycle is Solar Cycle 24, however, compared to previous cycles in the Modern Maximum, the solar activity during Cycle 24 was substantially lower. This is observable in the F10.7 solar index in Figure 2.4. The cycle before it, Cycle 23, is more representative of the level of activity seen in modern times and was therefore selected as the benchmark for obtaining low and high solar activity data.

![Figure 2.4: F10.7 solar index from January 1986 to January 2021, showing the four recent solar cycles 22, 23, 24 & 25. (Data obtained from the NASA Goddard Space Flight Center)](image)

<table>
<thead>
<tr>
<th>Date</th>
<th>Maximum 1st March 2000</th>
<th>Minimum 1st January 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily</td>
<td>215.1</td>
<td>78.7</td>
</tr>
<tr>
<td>81-day Average</td>
<td>181.8</td>
<td>74.2</td>
</tr>
</tbody>
</table>

1The Modern Maximum refers to a period of relatively high solar activity which began with Solar Cycle 15 in 1914
 CHAPTER 2. INITIAL EXPLORATION OF SATELLITE AERODYNAMICS

Solar cycles are typically defined by the sunspot count and the maximum sunspot count of Cycle 23 occurred in November of 2001 while the minimum count, which marks the boundary between Cycles 23 and 24, occurred in December 2008. While these are useful dates, they do not guarantee reliable values for the F10.7 which is required by the atmospheric models to predict the properties of the atmosphere. It is useful to find periods with relatively stable F10.7 as this avoids the effects of strong solar weather events influencing the properties of the atmosphere. For this thesis, the solar maximum and minimum were taken to be 1st March 2000 and 1st January 2008 respectively. The daily F10.7 solar index as well as the 81-day average centred on each date of the maximum and minimum are shown in Table 2.2.

2.3.3.2 Density of the Atmosphere

As seen in Figure 2.3, there is significant variation in the density of the atmosphere at a given altitude. In addition to the long-term variations brought about by the solar cycles, the density of the atmosphere is also affected by short term factors such as the underlying Earth topology as well as the day/night cycle. The graphs in Figure 2.5 show the density distribution of the atmosphere with altitude in the VLEO region using NRLMSISE-00, while the graphs in Figure 2.6 show the relative density distribution of the atmosphere with altitude for the same range. As can be seen from Figures 2.5a & 2.5b, which are for the solar maximum and minimum respectively, the average density of the atmosphere is generally higher during periods of solar maximum.

For the dates used here, the mean atmospheric density is about 3.5 times higher during solar maximum compared to the solar minimum at 250km but drops to only 1.3 times larger at 150km. The range of densities also appears to be much tighter during high solar activity than during low as can be seen from Figures 2.6a & 2.6b. For instance, at 200 km in altitude, 80% of the atmosphere has a density within ±10% of the mean density for that altitude during high solar activity, while the 10-90 range during low solar activity is much broader, at ±20%. It is worth noting that this spread may only be a feature of the dates shown here and not an overall trend of the atmospheric density. A comprehensive survey of the interquartile and 10-90 ranges of the atmosphere has not been performed here as this was beyond the scope of this project.

![Figure 2.5: Density distribution during Solar Maximum and Solar Minimum](image)

(a) Solar Maximum (March 2000)  
(b) Solar Minimum (January 2008)
2.3. THE ATMOSPHERE IN VLEO

2.3.3.3 Temperature of the Atmosphere

Another important property of the atmosphere is the temperature of the gas as this has a major effect on the simulation of rarefied flows. In particular, the temperature affects the thermal velocity of the particles in the fluid which changes the mean free path of the gas and in turn the Knudsen Number of the flow. From a simulation’s perspective, this can impact the predicted coefficients of drag for a given body. Figure 2.7 shows the variation of temperature with density across a range of altitudes in VLEO, during both high and low solar activity. Each cluster of points on the figure represents a single altitude with samples taken at 5° increments in both latitude and longitude. The dashed lines indicate the mean temperature and density for each altitude during high and low solar activity.

Figure 2.7: Variation of temperature with density across a range of altitudes in VLEO. Each cluster represents a single altitude during either high or low solar activity. The dashed line shows the mean temperature and density at each altitude during either high or low solar activity.
As can be seen from Figure 2.7, while the atmosphere is hotter and denser during high solar activity, there is no specific correlation between the density and the temperature. In particular, above 200km in altitude, there is little change in the mean temperature or range of temperatures experienced at each altitude for each solar activity level. Below 200km, the range of temperatures narrows and the mean temperature drops quite considerably as the atmosphere thickens and becomes more homogeneous.

For a given altitude the range of temperatures the satellite might experience is quite high. For instance, at 200 km the temperature range is 270K and 280K for high and low solar activity respectively. This variation could present difficulties when trying to choose representative conditions for calculating the $C_D$ values for orbital simulations in VLEO since even small variations in the $C_D$ could result in large changes in the orbit propagation and in turn the fuel demand.

### 2.3.3.4 Composition of the Atmosphere

The final key component of the gas is its molecular composition. NRLMSISE-00 models the relative composition of seven molecular species: Atomic Oxygen ($\text{O}_1$), Molecular Oxygen ($\text{O}_2$), Atomic Nitrogen ($\text{N}_1$), Molecular Nitrogen ($\text{N}_2$), Argon ($\text{Ar}$), Helium ($\text{He}$) and Hydrogen ($\text{H}$). Figure 2.8 shows the relative number fraction of each species with density, during solar maximum and solar minimum.

As can be seen from Figure 2.8, the two most dominant molecular species in VLEO are Atomic Oxygen (2.8a) and molecular Nitrogen (2.8b), which together comprise about 90% of the neutral atmosphere in this region. The other 10% is mostly made up of molecular Oxygen(2.8c), Helium (2.8g) and Atomic Nitrogen (2.8e). The least abundant are Argon and Hydrogen which make up less than 0.5% of the atmosphere.

The abundance of Atomic Oxygen is a major concern for VLEO operations. The Atomic Oxygen chemically reacts with exposed surfaces forming oxides that can be volatile or mechanically fragile. This can cause degradation of surfaces leading to a drop in optical, thermal or mechanical performance. Steps are often taken to mitigate these effects through careful material selection or by applying coatings where necessary[36]. The effects of Atomic Oxygen are explored in more detail in Chapter 5.

### 2.3.4 Reduced Atmospheric Model

As has been seen in the sections above, there is significant variability in the atmospheric properties of VLEO. For modelling and predicting the orbital paths in this region, it is essential to carefully include all these fluctuations. However, this is far too much detail for performing drag simulations. Nevertheless, the atmospheric properties used as part of these simulations must be representative of the gas at the altitude of interest. To achieve this, a reduced atmospheric model has been used based on data from NRLMSISE-00. This contains the global average values of the density, temperature and composition of the atmosphere for a range of altitudes in VLEO during both solar maximum and minimum. Figure 2.9 shows the density values for this reduced model, while the temperature and composition have been included in Figures 2.7 & 2.8 respectively. A summary of the properties at an altitude of 200km are presented in Table 2.3.
2.3. THE ATMOSPHERE IN VLEO

(a) Atomic Oxygen ($O_1$)  
(b) Nitrogen ($N_2$)  
(c) Oxygen ($O_2$)  
(d) Helium ($He$)  
(e) Atomic Nitrogen ($N_1$)  
(f) Argon ($Ar$)  
(g) Hydrogen ($H$)

Figure 2.8: Variation in the composition of key gas species in the Thermosphere across a range of altitudes in VLEO. Each cluster represents a single altitude during either high or low solar activity. The dashed line shows the mean composition and density at each altitude during either high or low solar activity.
Figure 2.9: Atmosphere density from NRLMSISE-00 used for Drag modelling

Table 2.3: Properties of the atmosphere at 200km on the 1st March 2000 at 00:00:00 UTC as given by the spherically weighted mean of NRLMSISE-00 model[30] using sampling points at 1° intervals in Latitude and Longitude

<table>
<thead>
<tr>
<th>Mission Profile</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude</td>
<td>200 km</td>
</tr>
<tr>
<td>Date</td>
<td>1st March 2000</td>
</tr>
<tr>
<td>Time</td>
<td>00:00:00 UTC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bulk Fluid</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number Density</td>
<td>$1.06 \times 10^{16} \text{ m}^{-3}$</td>
</tr>
<tr>
<td>Temperature</td>
<td>1063 K</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Composition</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Atomic Oxygen ($O_1$)</td>
<td>56.195 %</td>
</tr>
<tr>
<td>Nitrogen ($N_2$)</td>
<td>39.893 %</td>
</tr>
<tr>
<td>Oxygen ($O_2$)</td>
<td>2.972 %</td>
</tr>
<tr>
<td>Helium ($He$)</td>
<td>0.585 %</td>
</tr>
<tr>
<td>Atomic Nitrogen ($N_1$)</td>
<td>0.278 %</td>
</tr>
<tr>
<td>Argon ($Ar$)</td>
<td>0.064 %</td>
</tr>
<tr>
<td>Hydrogen ($H$)</td>
<td>0.013 %</td>
</tr>
</tbody>
</table>
2.4 Modelling Aerodynamic Forces in VLEO

2.4.1 Overview

The aerodynamic forces that a body experiences are caused by its interaction with the fluid medium through which it is passing. In this instance, the satellite body is interacting with the rarefied gas of the atmosphere in VLEO. To determine these forces, it is necessary to implement a flow solver, but it is important to choose an appropriate flow solver. This section will discuss the properties of the fluid in VLEO, the methods that are available for determining the aerodynamic forces in this medium and implementations of these methods that will be used later in the thesis.

2.4.2 Molecular Flow

While the density is higher in VLEO than in more traditional LEO, the density is still very low when compared to the atmospheric density at sea level as can be seen from Figure 2.3. A way of characterizing the low density flows in VLEO is to use the non-dimensional Knudsen Number ($K_n$), given by Equation 2.1, which describes the ratio between the molecular mean free path ($\lambda_0$), and the size of the object ($l_{ref}$).

$$K_n = \frac{\lambda_0}{l_{ref}} \quad (2.1)$$

$\lambda_0$ describes the average distance that particles travel between collisions. Assuming a Variable Hard Sphere (VHS) model, $\lambda_0$ for a gas mixture can be calculated using Equation 2.2, where $n$ is the number density of the gas and $n_p$ and $(\lambda_p)_0$ are the number density and mean free path of a molecular species $p$ respectively. The mean free path of a particular molecular species can be calculated using Equation 2.3, where ($d_{ref}$)$_{pq}$ is the average hardshell diameter of the molecular species $p$ and $q$ at the reference Temperature ($T_{ref}$)$_pq$ in Kelvin, $T$ is the temperature of the mixture in Kelvin, $\omega_{pq}$ is the viscosity-temperature index and $m$ is the mass of a single molecule of the species. Typical values for these variables are given in Table 2.4.

Table 2.4: Molecular Species data required to implement the VHS model and used to calculate $\lambda_0$.

<table>
<thead>
<tr>
<th>Gas Species$^1$</th>
<th>Molecular Mass [kg]</th>
<th>Molecular diameter [m]</th>
<th>Viscosity-temperature index ($\omega$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atomic Oxygen ($O_1$)</td>
<td>$2.65 \times 10^{-26}$</td>
<td>$3.00 \times 10^{-10}$</td>
<td>0.80</td>
</tr>
<tr>
<td>Nitrogen ($N_2$)</td>
<td>$4.65 \times 10^{-26}$</td>
<td>$4.07 \times 10^{-10}$</td>
<td>0.74</td>
</tr>
<tr>
<td>Oxygen ($O_2$)</td>
<td>$5.31 \times 10^{-26}$</td>
<td>$3.96 \times 10^{-10}$</td>
<td>0.77</td>
</tr>
<tr>
<td>Helium ($He$)</td>
<td>$6.65 \times 10^{-27}$</td>
<td>$2.30 \times 10^{-10}$</td>
<td>0.66</td>
</tr>
<tr>
<td>Atomic Nitrogen ($N_1$)</td>
<td>$2.33 \times 10^{-26}$</td>
<td>$3.00 \times 10^{-10}$</td>
<td>0.80</td>
</tr>
<tr>
<td>Argon ($Ar$)</td>
<td>$6.63 \times 10^{-26}$</td>
<td>$4.11 \times 10^{-10}$</td>
<td>0.81</td>
</tr>
<tr>
<td>Hydrogen ($H$)</td>
<td>$3.34 \times 10^{-27}$</td>
<td>$2.88 \times 10^{-10}$</td>
<td>0.67</td>
</tr>
</tbody>
</table>

$^1$ Species in order of abundance at 200km
$^2$ At the reference temperature 273.15 K ($T_{ref}$)
\[ \lambda_0 = \sum_{p=1}^{s} \left( \frac{n_p}{n} \right) (\lambda_p)_0 \]  
(2.2)

\[ (\lambda_p)_0 = \frac{1}{\sum_{q=1}^{s} \left( \pi \left( d_{ref} \right)^2 p_q n_q \left( \frac{T_{ref}}{T} \right)^{\omega_{pq}^{-1} \left( 1 + \frac{m_p}{m_q} \right)^{1/2}} \right)} \]  
(2.3)

The mean free paths for a number of altitudes during high solar activity are shown in Table 2.5. These were calculated using the reduced atmospheric model described in Section 2.3.4 and Equations 2.1-2.3. Due to the low density, the fluid medium has comparatively high mean free paths, on the order of 10-100s of metres. Assuming the length of the satellite is used as the reference length (which is normally <10m) the \( K_n \) of the fluid in VLEO is quite large. For example, the \( K_n \) for bodies that are 0.5m in length varied from 48.8 at 150km to 2176.1 at 300km, as seen in Table 2.5. Fluids with high Knudsen Number (\( K_n >> 1 \)) such as these are generally described as ‘free molecular flow’ [46]. So, this would suggest that the satellites experience free molecular flow right down to 150 km.

Table 2.5: Variation of Mean Free Path (\( \lambda_0 \)) and Knudsen Number (\( K_n \)) with altitude for satellite with a reference length of 0.5m, Number Density from the last Solar Maximum (March 2000)

<table>
<thead>
<tr>
<th>Altitude [km]</th>
<th>Number Density [1/m^3]</th>
<th>Mean free path [m]</th>
<th>Knudsen Number [-]</th>
</tr>
</thead>
<tbody>
<tr>
<td>150</td>
<td>5.470 \times 10^{16}</td>
<td>24.4</td>
<td>48.8</td>
</tr>
<tr>
<td>200</td>
<td>1.081 \times 10^{16}</td>
<td>130.7</td>
<td>261.5</td>
</tr>
<tr>
<td>250</td>
<td>3.663 \times 10^{15}</td>
<td>417.8</td>
<td>835.6</td>
</tr>
<tr>
<td>300</td>
<td>1.496 \times 10^{15}</td>
<td>1088.0</td>
<td>2176.1</td>
</tr>
</tbody>
</table>

1 Reference length = 0.5m

There are also local effects to consider, such as the accumulation of particles around the front of the satellite caused by the slower velocity of particles reemitted from the surfaces following a diffuse interaction. These particles linger around the front of the satellite creating a region of higher density which raises the probability of particle-particle collisions taking place. This was observed during initial simulations of blunt test bodies using DSMC methods. Under the atmospheric conditions at 200km during high solar activity, the local \( K_n \) dropped from 261.5 in the far-field to as low as 2 in the region immediately ahead of the test body. There also appeared to be a degree of variability in this reduction, with the minimum having some dependence on the geometry. For instance, the lowest observed \( \lambda_0 \) rose to about 4.5 metres when testing sharp conic profiles. Since \( K_n \) is close to unity the flow in the region ahead and around the satellite could be considered to be transitional rather than fully free-molecular [46].

2.4.3 Methods for Determining Forces in a Molecular Flow

For flows with high \( K_n \), i.e. ‘molecular flows’, continuous methods such as Navier-Stokes are not valid for simulating the fluid. There are several alternative methods available for modelling the interaction between the gas and the satellite, and these can broadly be grouped into Panel Methods, Ray-Tracing Panel (RTP), Test-Particle Monte Carlo (TPMC) and DSMC.
2.4. MODELLING AERODYNAMIC FORCES IN VLEO

2.4.3.1 Panel Methods

In this method, the satellite body is divided into discrete face elements (or panels) and the interactions of gas particles are modelled for each face element individually. This is achieved by treating the incoming fluid as a collimated flow, accounting for the thermal variation with that column using a kinematic equilibrium function, such as the Maxwell-Boltzmann distribution (This is discussed in more detail in Section 5.3). The key assumption here is that in a free molecular flow the particle-particle interactions within the fluid are negligible which means that particles that reflect off the body do not affect the upstream flow. The forces on the surface element can therefore be determined using a Gas-Surface Interaction Model (GSIM) based purely on the properties of the free stream. The total forces on the body can then be calculated by adding up the forces on each element.

The key advantage of these types of models is that they are very fast when compared to other 'molecular flows' solvers, since the speed of computation is primarily determined by the number of discrete panels. However, they also have some key limitations. For instance, most panel methods are not capable of dealing with faces that are totally or partially shaded by another surface, as the method assumes that all panels may be interacted with regardless of orientation or upstream obstructions. While there may be some interactions with these surfaces due to the random thermal motion of some of the particles, the level and distribution of these interactions are dependent on the upstream geometry. Panel Methods are also not capable of accurately determining the forces on concave bodies as particle reflections are not accounted for. A concave body has one or more sections that curve inwards or are recessed, allowing particle reflections from one part of the body to strike another part.

2.4.3.2 Ray-Tracing Panel (RTP)

As the name suggests, RTP employs ray-tracing techniques, that were originally developed and refined for 3D graphics processing, to track the progress of the fluid around the satellite geometry. This is achieved by once again assuming the incoming fluid can be treated as a collimated flow, which is analogous to a beam of light. This has certain advantages over Panel Methods as ray tracing natively accommodates the shading of faces by upstream obstructions. Additionally, by implementing a suitable GSIM, particle reflections can also be simulated. This allows RTP to be applicable to more complex concave geometries, as may be seen on modern satellites.

The primary limitation of this method is choosing a suitable and representative number of rays for the incoming fluid. Keeping with the graphics processing analogy, this is equivalent to choosing the resolution of the image, so the more rays you include, the better the approximation of the forces on the satellite body. In contrast to panel methods, the processing time scales with the number of rays used. Additionally, RTP works best when the reflections implemented by the GSIM are specular or quasi-specular. As will be discussed in Section 5.3, diffuse reflections result in the dispersal of the beam which if tracked can be computationally intensive.
2.4.3.3 Test-Particle Monte Carlo (TPMC)

The TPMC method was originally developed by D. Davis in 1960[47]. In this method, the fluid is modelled as a set of ‘test-particles’, where each simulated particle represented a large number of real molecules with similar properties. The velocities of these simulated particles are composed of the free stream velocity and the thermal velocity which is determined probabilistically using a kinematic equilibrium function, such as the Maxwell-Boltzmann distribution which allows for a degree of dispersion in the fluid. These particles can interact with, and reflect off surfaces of the simulated geometry which, much like RTP, allows TPMC to be applicable to more complex concave geometries. However, these simulated particles do not interact with one another following the general assumption of free molecular flow. The key drawback of this method is that the computation time scales with the number of simulated particles.

2.4.3.4 Direct Simulation Monte Carlo (DSMC)

The DSMC method is a probabilistic physical simulation of a fluid, first pioneered by G. Bird in 1976[48, 45]. The method directly simulates the motion of each particle (or a statistically significant group of particles) within the fluid. As with TPMC, particle-surface interactions are determined by a GSIM and allow for multiple reflections with the simulated body. However, unlike TPMC, simulated particles in the DSMC method can also interact with other particles in the fluid. These particle-particle interactions are probabilistic and can be simulated using a molecular collision model such as the Hard Sphere (HS) model, Variable Hard Sphere (VHS) model, or the Variable Soft Sphere (VSS) model.

Since this method captures more of the particle interactions, it generally has improved accuracy over some of the analytical models described above. Several studies have shown good agreement between the results of simulations performed by DSMC models and observations of low flying satellites and re-entries[42, 50]. However, performing simulations using DSMC methods can be time-intensive and computationally expensive because the simulated flow needs time to converge to a steady-state. For regions in high VLEO where the density is low and the collisions are infrequent, the convergence time is relatively quick, however, it presents a challenge lower in the atmosphere where the density is high, particularly for large three-dimensional bodies. See Appendix A.1 for discussion on the conversion rate of the software tool SPARTA.

2.4.3.5 Method Choice and available Tools

Of the four methods described above, DSMC was chosen as the primary method for obtaining the aerodynamic forces in this thesis. While panel methods and RTP are useful for establishing an estimate of the drag, it is desirable to account for all aspects of the geometric effects including any non-linearities. This includes reflections and the increased particle-particle encounters in the higher density regions around the satellite (of which the latter is not accounted for in TPMC).

Since first being developed in 1976 by G. Bird[48, 45], there have been a variety of implementations of the DSMC method by various groups from academia to industry. Some of the DSMC tools that are available at the time of writing are summarized in Table 2.6. DS1V, DS2V & DS3V are...
the original implementations of the DSMC method developed by Bird, in 1, 2 and 3 dimensions respectively. These are well-established codes with many citations and are often considered the gold standard in DSMC. However, the underlying code is now quite old and not capable of supporting multi-threading which limits its speeds compared to newer codes.

Examples of more modern implementations include MONACO and DAC developed by Cornell University, and the Johnson Space and Langley Research Centers respectively. Both implementations have been developed and optimized for parallel computing and have been used by NASA on several projects including Mars entry and descent. Unfortunately, these codes are only available to US persons and institutions.

There are also many open-source DSMC codes available, including dsmcfoam and SPARTA. dsmcfoam is an implementation of the DSMC method within the framework of OpenFOAM (which stands for Open-source Field Operation And Manipulation), an open-source Computational Fluid Dynamics (CFD) tool-kit supported by OpenCFD. It boasts friendly syntax and a wide range of ready to use applications and models, however, it also has a fragmented developer community with multiple forked projects.

The DSMC tool used for the work in this thesis was SPARTA developed by Sandia National Laboratories[51]. It is also an open-source tool that offers both two-dimensional and three-dimensional capability and has been tested on other spacecraft [52, 53]. It has a steep learning curve but the code is generally well documented.

As will be seen in Chapter 3, Panel Methods will be used as a stand-in for DSMC to test the framework. Analytical methods such as these are fast and predictable, which makes developing complex frameworks easier. The panel method implementation used in this thesis was developed by L. Sentman in 1961 [54].

2.4.4 Gas-Surface Interaction Model

An important aspect of modelling rarefied gases is the interaction between the particles and the surfaces of the test body. Both the Sentman method and the DSMC tool SPARTA use a diffuse model to replicate the particle-surface interactions[54, 56] within a rarefied gas. In this diffuse model, interactions are governed by the surface temperature and the energy accommodation coefficient [48]. The energy accommodation coefficient $\alpha_T$ describes how much energy the incoming particle loses to the surface and is defined as

$$\alpha_T = \frac{E_i - E_r}{E_i - E_w}$$  (2.4)

where $E_i$ is the kinetic energy of the incident molecule, $E_r$ is the kinetic energy of the re-emitted molecule, and $E_w$ is the kinetic energy the re-emitted molecule would have if it were re-emitted from the surface at the surface’s temperature. When $\alpha_T$ equals 1, complete thermal accommodation has occurred, i.e. the particle has been re-emitted at the same temperature as the impacted surface and therefore has lost some energy to the surface.
### Table 2.6: A Qualitative comparison of Molecular Flow solvers

<table>
<thead>
<tr>
<th>Method/Software</th>
<th>Type</th>
<th>License</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentman Method</td>
<td>Panel</td>
<td>unclassified</td>
<td>This is a panel method based on the work by Lee H. Sentman of Lockheed Missiles and Space. It uses a statistical model of the gas to simulate the variation in molecular velocities.</td>
</tr>
<tr>
<td>DS2V/DS3V</td>
<td>DSMC (parallel)</td>
<td>closed-source freeware</td>
<td>DS2V/DS3V was first developed by Prof. Bird around 1963 and is a well-established visually interactive DSMC program for steady or unsteady flows.</td>
</tr>
<tr>
<td>OpenFoam</td>
<td>DSMC (parallel)</td>
<td>open-source GPL-v3</td>
<td>dsmcFoam is an open-source DSMC code built within the framework of the toolbox OpenFOAM supported through OpenCFD Ltd.</td>
</tr>
<tr>
<td>SPARTA</td>
<td>DSMC (parallel)</td>
<td>open-source GPL-v3</td>
<td>Stochastic PArallel Rarefied-gas Time-accurate Analyzer (SPARTA) is a DSMC code created by the Sandia National Laboratories.</td>
</tr>
<tr>
<td>DAC</td>
<td>DSMC (parallel)</td>
<td>freeware U.S. persons only.</td>
<td>The DSMC Analysis Code (DAC) is a Rarefied gas modeling software developed by Lawrence Research Centers and NASA Glenn Research Center.</td>
</tr>
</tbody>
</table>

#### Advantages
- Computationally quick compared to DSMC methods (therefore good for getting an approximate order of magnitude)
- Appears to be shape independent, working well for all shapes and collisional models

#### Disadvantages
- No initialization/convergence criteria
- Boundary conditions are not documented
- Old software
- Not optimized for parallel computing
- Mutually inconsistent and collisional models
At present, there is some uncertainty around the $\alpha_T$ of materials in VLEO. Empirical findings from satellites in orbit have shown that $\alpha_T$ has a value in the range of 0.86 to 1, with limited dependences on the underlying material\cite{27,32}. This contrasts, however, with early ground testing, which has shown some dependence on the material choice, though often these are not comparable conditions to those seen in orbit. While it is likely that many materials will still exhibit mostly diffuse or quasi specular behaviour, this is not known with any great certainty. Recent work by Roberts et al.\cite{22,33}, seeks to improve this understanding by examining the gas-surface interactions between the material and the atmosphere under the conditions observed in orbit. At the time of writing, the Rarefied Orbital Aerodynamics Research facility (ROAR) is still being commissioned, but once operational will be capable of simulating the free molecular flow and levels of atomic oxygen flux observed in VLEO.

For the work presented here an $\alpha_T$ of 0.95 was selected, representing a mostly diffuse environment with some specular reflections occurring. The surface temperature of the satellites is taken to be 300K, which is broadly within the operational range of most terrestrial satellites. These values were used to define the properties of the surface for both the Panel Methods and the DSMC simulations.
2.5 Result & Discussion

2.5.1 Drag and Fuel

This section presents the results of aerodynamic calculations performed on the candidate profiles described in Section 2.2.2. The calculations were performed assuming an altitude of 200 km during the solar maximum (Section 2.3.3.1) using the DSMC tool SPARTA (Section 2.4.3.5). Also presented are estimates of the required fuel for a generic propulsive system.

Table 2.7 shows the aerodynamic forces the four profiles would experience at an orbital altitude of 200 km during high solar activity. Also shown in the table are predicted fuel requirements for each profile. To calculate these fuel requirements, it is assumed that the thrust produced by the propulsion system matches the drag experienced. Then at any given time, the fuel mass flow \( \dot{m}_f \) can be calculated using Equation 2.5, where \( F_T \) is the thrust and \( I_{sp} \) is the specific impulse of the propulsion system. As seen in Section 2.3.3 there is a significant amount of variability in the density of the atmosphere at a given altitude, however, this has been accounted for in the reduced atmospheric model used to calculate the atmospheric properties. This means the results should be broadly representative of the average fuel flow experienced at 200km altitude during high solar activity. Subsequently, a rough extrapolation for mission length \( T_m \) can be made using Equation 2.6, assuming the level of solar activity remains broadly the same.

\[
\dot{m}_f = \frac{F_T}{g_0 I_{sp}} \quad (2.5)
\]

\[
m_f = \dot{m}_f T_m \quad (2.6)
\]

Table 2.7 shows the aerodynamic forces the profiles would experience at 200km during high solar activity. Also shown in the table are predicted fuel requirements for the profile assuming constant drag and matched thrust. It was seen that in general tapering the nose of the profile did improve the drag on the body of the satellite. In a 200km altitude orbit, the drag on a satellite with profile S4 (a 0.5m taper) is calculated to be 8.88mN compared to 9.63mN for an S1 profile which has no taper. This would correspond to an 8% reduction in the fuel required to maintain altitude (Table 2.7).

Table 2.7: Comparison of Aerodynamic Forces for each profile at 200km as well as estimated fuel requirements. Ref. values: \( v = 7,794 m/s, \rho = 5.30521 \times 10^{-10} kg/m^3, A_{ref} = 0.25 m^2 \)

<table>
<thead>
<tr>
<th>Profile</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forces</td>
<td>Drag ([mN])</td>
<td>9.63</td>
<td>9.32</td>
<td>9.16</td>
</tr>
<tr>
<td></td>
<td>(C_D) ([-])</td>
<td>2.391</td>
<td>2.314</td>
<td>2.273</td>
</tr>
<tr>
<td>Fuel rate for (I_{sp}) ([\times 10^{-7}kg/s])</td>
<td>1000 s</td>
<td>9.82</td>
<td>9.50</td>
<td>9.33</td>
</tr>
<tr>
<td></td>
<td>2000 s</td>
<td>4.91</td>
<td>4.75</td>
<td>4.67</td>
</tr>
<tr>
<td></td>
<td>3000 s</td>
<td>3.27</td>
<td>3.17</td>
<td>3.11</td>
</tr>
</tbody>
</table>
It should be noted that the forces in Table 2.7 were calculated using globally averaged properties for the atmosphere. However, as seen in Section 2.3, the density of the atmosphere varies considerably across the globe, so for a given orbit the actual drag force experienced may vary by $\pm 10\% - 20\%$ with respect to this value. Ideally, in a perfect drag compensation scheme, the thrust would match the drag experienced as close as possible, as in the case of GOCE. However, GOCE’s payload was a highly sensitive gradiometer that was tied into the drag compensation system. Practical systems may not have access to accelerometers with sufficient sensitivity to measure the changes in drag force needed to achieve perfect drag compensation. This will result in a degree of over and under thrusting over the course of the orbit which may affect the efficiency of the platform and thus the predicted fuel requirement.

2.5.2 Sources of Drag

In the previous section, it was shown that the improvements in drag can be achieved by tapering the nose of the profile. It is useful to know how the surfaces of the profiles contribute to the overall forces that are experienced to help identify areas of further improvement.

Figure 2.10 shows the total force and force contributions due to components of the shear and pressure forces on the satellite as calculated using DSMC. In this instance pressure specifically refers to the normal force on a face while the shear is a force parallel to the surface of that face. Figure 2.11 shows the distribution of the pressure and shear stress along the surfaces of the profiles that contribute to the drag these bodies experience. In all cases, the origin of the surface is at the lower rear (right) corner of the profile as shown in Figure 2.12. Figure 2.12 also provides a schematic view of the forces seen in Figure 2.11 on a tapered and un-tapered profile, however, these forces are not to scale. Finally, Figure 2.13 shows the normalised density field around a tapered and un-tapered profile. The density field is normalised to the density of the far-field as given in Table 2.3 with the dotted red line indicates regions of the density field that are equal to the far-field density.

![Figure 2.10: Comparison of Drag Coefficients for the profiles at 200 km](image-url)
CHAPTER 2. INITIAL EXPLORATION OF SATELLITE AERODYNAMICS

(a) Profile S1

(b) Profile S2

(c) Profile S3

(d) Profile S4

Figure 2.11: Distribution of pressure and shear stress along the surfaces of Profiles S1 and S3 that contribute to the drag of the profile.

For the purpose of performing generalised comparisons with the un-tapered profile S1, the tapered profiles will be represented by profile S3 as seen in Figures 2.12b & 2.13b. Profile S3 was chosen as representative of the tapered profiles because it had a smaller internal nose angle than profile S2’s 45° nose angle and a shorter nose length than profile S4, thus providing a middle ground between the other two tapered profiles.

As would be expected, surfaces facing directly into the flow (including the tapered sections of profile S3, surfaces 2 and 4), are the predominant contributors to the drag the profile experience as these surfaces experience higher particle fluency than surface parallel to the flow. The surfaces parallel to the flow do contribute to the drag, however, the shear stresses along these surfaces is 10-20 times smaller than the normal pressure on the front-facing surface. Figures 2.13a & 2.13b clearly show the increase in density around the front of the two profiles as the particles pile up. For the sake of clarity, the colour map and contours have been truncated to 10 times the ambient atmospheric density. Ahead of the un-tapered profile S1, the density peaks at 27 times the ambient density indicating a region of more frequent particle interaction. In contrast, near the centre of the nose of profile S3 the density peaks at 24 times the ambient density and the peak density is only 15 times higher along the tapered surfaces.

32
Tapering the forward surfaces does appear to reduce the drag, as was observed in the previous section. On the tapered sections, these particles will be glancing the surface rather than striking ‘head-on’. Glancing blows have a lower normal velocity than head-on collisions so if the particle-surface interaction were perfectly specular, the force vectors would be aligned with the normal vector of the surface. By balancing this vector with a surface on the other side of the body to cancel out the lateral forces, a smaller force is left in the velocity direction. If the profile were asymmetrical then the body would experience a small lateral force. However, while the pressure drag decreases with increased taper, there is a large increase in the overall shear drag, as seen in Figure 2.10. The GSIM used here assumed an energy accommodation coefficient of 0.95, implying particles achieve close to full thermal accommodation with the surfaces and thus experience predominately diffuse interactions. The diffuse interaction between the particles and the surfaces introduces shear stress along these tapered sections. Nevertheless, despite the increased area exposed to shear forces, tapering the nose still reduces drag on the body overall.

As can be seen from Table 2.1, all of the tapered profiles tested had a blunted nose with a width of 0.1m. This was included to be consistent with previous work carried out in this area. However, as can be seen from Figure 2.11c, the pressure on this surface (panel 3 of Profile S3) is similar to the pressure experienced on the front of the cuboid body Profile S1 (panel 2). So, from a pure drag perspective, including a blunted nose would not be beneficial. However, it may help to balance volume requirements with minimizing drag, so this would need to be explored further.

As noted earlier, the shear stresses along the surfaces parallel to the flow are 10-20 times smaller than the normal pressure on the front-facing surfaces of both profiles. While these shear stresses are not large, these surfaces contribute about 10% of the overall drag. It, therefore, raises the question of whether these surfaces can be shaped in some way to reduce their contribution to the drag. One way might be to taper these rear surfaces, thereby turning them away from the incoming flow. In general, surfaces that are pointing away from the incoming flow are shielded by the body and therefore experience limited particle flux and thus negligible force. This is observed for panels 4 and 6 on Profiles S1 and S3 respectively, as can be seen in Figure 2.11. This is also clearly seen from the density fields in Figure 2.13, which shows a region of low density directly behind the satellite body.
CHAPTER 2. INITIAL EXPLORATION OF SATELLITE AERODYNAMICS

Figure 2.13: Normalised density field around the profiles. Density field normalised to the free-stream density (red dotted line)
2.5.3 Results & Discussion: Software Performance Comparison

As is discussed in Section 2.4 there are several software tools available for simulating free molecular flow. One of the tools considered is the DSMC method SPARTA. This is an open-source code and while it is well cited it is worth confirming that it is a reliable implementation of the method. To do this, the aerodynamic forces calculated by SPARTA will be compared to results by DS2v, often considered the gold standard in DSMC codes and the panel method based on the work by Sentman[54].

Figure 2.14 and Table 2.8 show the coefficient of drag experienced by each of the profiles described in Table 2.1 calculated using either DS2V, SPARTA in 2D, SPARTA in 3D or the Sentman Method, at an altitude of 200km. Table 2.8 also shows the relative difference to DS2V. As can be seen from Figure 2.14 and Table 2.8 DS2V and Sparta 2D generally correlate well with one another with an error of less than 3%. Both codes are stochastic methods, thus there will be some random variation in the results but the most likely reason for the difference in the results is the implementation of the code and handling of the surface boundary conditions.

By contrast, there is a consistently higher difference between the Sentman method and either of the 2D DSMC methods across most of the profiles tested of between 3-5%. The satellite bodies described in Section 2.2.2 are 2D convex shells that do not experience secondary particle impacts from prior particle-surface interactions. This would usually suggest that panel methods would be sufficient for determining the aerodynamic forces on these bodies in a rarefied gas. However, the difference in the results shown here could be a result of ‘particle pileup’ ahead of the test bodies which is more completely captured in the DSMC methods but not normally considered in panel methods.

Under the atmospheric conditions at 200km during high solar activity, the mean free path of the fluid for these simulations is 131m (\(K_n = 262\)). It was observed that ahead of the S1 profile, the local mean free path of the fluid dropped to 1m immediately ahead of the profile (\(K_n = 2\)), in both the SPARTA and DS2V simulation. This increased to 4.5m for the sharper S4 profile (\(K_n = 9\)). Since \(K_n\) is close to unity the flow in the region ahead and around the satellite could be considered to be transitional rather than fully free-molecular[46].

Table 2.8: Comparison of Drag Coefficients from simulations performed using DS2V, Sentman and Sparta 2D & 3D with errors relative to DS2V results

<table>
<thead>
<tr>
<th>Profile</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drag Coefficient</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DS2V</td>
<td>2.391</td>
<td>2.314</td>
<td>2.273</td>
<td>2.204</td>
</tr>
<tr>
<td>Sentman</td>
<td>2.500</td>
<td>2.407</td>
<td>2.360</td>
<td>2.276</td>
</tr>
<tr>
<td>Sparta 2D</td>
<td>2.369</td>
<td>2.300</td>
<td>2.233</td>
<td>2.141</td>
</tr>
<tr>
<td>Sparta 3D</td>
<td>2.704</td>
<td>2.534</td>
<td>2.454</td>
<td>2.342</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>% Difference to DS2V</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentman</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Sparta 2D</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Sparta 3D</td>
<td>13</td>
<td>9</td>
<td>8</td>
<td>6</td>
</tr>
</tbody>
</table>
Sparta 3D was included as it was being considered for use in future work. It can be seen from both Table 2.8 and Figure 2.14 that the results of Sparta 3D are consistently higher. The source of this higher drag can simply be attributed to the shear drag along the two additional faces on either side of the body which are not simulated in two dimensions.

An important consideration is the time taken to perform the simulations. Panel methods, such as the Sentman method used here, are generally very quick as the computational complexity normally scales with the number of surfaces of the geometry. The candidate profiles used here have a maximum of 6 surfaces, so the simulation time was less than a second per profile. DSMC simulations are computationally more complex as they simulate the particle physics of the rarefied gas around the satellite with the complexity scaling with the number of particles simulated. Some DSMC implementations, like SPARTA, use a reduced number of simulated particles to represent the molecules in the gas. In this case, one simulated particle represented $1^{13}$ real particles. For the simulations performed here, it took around 1 hour for the simulation to converge to an uncertainty in the drag force of less than $1 \times 10^{-3}$ (when run on a single processor core). The trade-off between run time, simulated particle ratio and accuracy is discussed further in Appendix A.1.

The long simulation time is a major concern and it should be noted that the analysis performed as part of this section features results from only a handful of situations as a direct result of the long simulation time. This would normally make it prohibitive to explore an entire design space with sufficient fidelity. As described by Forrester et al. [58], this problem can be effectively managed by employing a surrogate model framework. This reduces the number of simulations required through careful selection of the sample points and effective interpolation of the data.

![Figure 2.14](image-url)

**Figure 2.14**: Comparison of Drag Coefficients from simulations performed using DS2V, Sentman and Sparta 2D & 3D.
2.6 Summary

Operating in VLEO presents some challenges, most notably the increased atmospheric drag due to the lower altitude. This chapter has presented an initial exploration of the atmospheric drag experienced by satellites in VLEO and, in particular how, the geometry of an aeroshell profile can affect the drag experienced by microsatellites in VLEO. It was demonstrated that a tapering the nose section can help to reduce the drag on a satellite in VLEO and, in turn, reduce the fuel required by the platform. However, it was also observed that, for the candidate profiles examined, this came at the expense of internal volume which can limit the profiles’ usability. It was therefore identified that it would be beneficial to explore the interaction between volume and drag for the aeroshell profile.

It was previously noted that it is computationally expensive to perform the simulations using DSMC methods, as a result of the convergence time of the particle stream. Despite this, it is still desirable to use DSMC methods for examining the geometry in order to account for non-linearities, such as particle reflections and the increased particle-particle encounters in the higher density regions around the satellite, which are not captured by faster panel methods. A potential mitigation is to approximate the results using a surrogate model, which would reduce the number of simulations required through careful selection of the sample points and effective interpolation of the data. This will be explored further in Chapter 3.
CHAPTER 3

GENERATION AND VALIDATION OF A SURROGATE MODEL FOR RAREFIED GAS MODELLING

3.1 Introduction

In Chapter 2 it was identified that it would be desirable to use DSMC methods to optimize the geometry of a satellite’s central body. However, DSMC methods, like other advanced flow solvers, are computationally heavy simulations that require a lot of processing time per profile candidate. It was therefore also determined that to fully and efficiently explore the satellite design space using a DSMC-based methodology a surrogate model would be needed. As outlined by Forrester et al. [59], a surrogate model approximates the ‘black box’ function using carefully selected observations of the original function and implementing effective interpolation between these sites. By carefully selecting and tuning the components of the surrogate model it should be possible to minimize the number of samples that are needed to fully explore a given design space for the central body.

There are three primary parts to a successful surrogate model: the sampling strategy, the interpolation method and the ‘black box’ function to be approximated. There are also secondary components that operate in a supporting role and ensure the successful generation of the model. This chapter presents the current state of surrogate model development and describes the components implemented as part of this research. It will also demonstrate how these components were tested and validated.
CHAPTER 3. GENERATION AND VALIDATION OF A SURROGATE MODEL FOR RAREFIED GAS MODELLING

3.2 Background

Computer-based simulation and analysis such as those based on DSMC are used extensively in engineering for a variety of tasks. However, despite the steady and continuing growth of computing power and speed, the computational cost of complex high-fidelity engineering analyses and simulations can limit their use in important areas like design optimization. Much of the current development in surrogate-based modelling builds on the framework developed by Sacks et al. [60] from 1989. Since then there have been many advances in the field as is reported in the review papers by Queipo et al. [61] and Simpson et al. [62]. More recently, Forrester et al. [59] presented an in-depth look at more recent surrogate interpolation and infill methods. These include some of the more popular interpolation methods such as Polynomial Response Surface Model (RSM), RBF and Kriging which are discussed below.

One of the oldest and perhaps simplest forms of surrogate model interpolation is the Polynomial Response Surface Model (RSM). RSM shares a similar methodology to polynomial regression for 1 dimension variables, whereby the model function is taken to be a sum of monomial components whose coefficients are calculated by means of least squares [63]. RSM's simplicity and low cost, when compared to some of the other interpolation methods described below, make it an ideal method for identifying the trend of the data quickly. However, RSM surrogates are generally considered unsuitable for non-linear and higher multi-dimensional problems [59], such as those that are of interest to this area of study. While these problems can be alleviated by limiting the size of the domain and the number of variables, thus effectively linearizing the problem, this would limit the effectiveness of performing global optimization on the design space. Therefore, while this framework may be useful to develop an initial understanding of the trend of the data, RSM does not present an appropriate approach for the work to be pursued.

A commonly used interpolation method for surrogate models is the Radial Basis Function (RBF) [64]. RBF is a type of Kernel smoothing technique that uses the weighted sum of simple functions (the kernel or basis function) to approximate the 'black box' function of the design space. In the case of RBF, the basis functions have radial components measured from the observed sample. The output of the basis function varies with the distance from the training point. The choice of kernel function is important and many examples of potential basis functions exist, including linear and polynomial, as well as more complex functions such as Gaussian and multi-quadratic. The weightings of the kernel functions are then calculated based on a set of training data, which generates an approximation of the 'black box' function. Depending on the choice of basis function, an RBF is highly tunable which can improve the approximation of the 'black box' function. However, this can also make the surrogate model unstable, so part of this chapter will address how the RBF is constructed, how it was tuned and what additional measures were needed to ensure an accurate estimation.

Kriging [65] is a popular and specific type of RBF which uses a basis function similar to a Gaussian. Kriging is named for Danie Krige who developed the method for geological surveying and mining during the 1950s [65]. This method was further expanded upon and introduced to engineering applications by Sacks et al. [60] who applied it to computational experiments. The method is now
very well developed and well understood, with many resources and explanations of the method available [66, 58].

As stated above, Kriging and RBF interpolation methods share very similar model formulations. Simpson et al. [67] demonstrated that RBF and Kriging perform well over a range of sampling strategies and sample sizes when compared to other interpolation methods such as RSM. For this research, RBF interpolation was selected as the most suitable interpolation method based on the work referenced above and on the recommendation of several comparison studies [67, 68, 69, 70]. As will be seen in Chapter 4, the number of input parameters needed to define a satellite’s geometry can vary significantly depending on the type of profile. Since RBF scales well with the number of input parameters [64], RBF is ideally suited for this application.

An important aspect of generating a surrogate model is the sampling of the design space. Review papers such as those by Giunta et al. [71], Koehler and Owen [72], Simpson et al. [67], and Chaloner and Verdinelli [73] provide a detailed overview of the common space-filling methods. Generally, these can be divided into a number of categories including single-stage methods (such as the Fully Factorial and Latin Hypercube Sampling), sequential sampling methods without adaption to the data and sequential with adaption to the data.

The simplest single-stage method is using an orthogonal grid of sampling points often referred to as Fully Factorial sampling. As demonstrated by Simpson et al. [67] these methods perform well at returning low max error values, particularly as they enforce the placement of samples at the corners of the domain. Though it was also noted that Fully Factorial may not provide good overall coverage when compared to other sampling methods.

A popular approach for achieving space-filling is the Latin Hypercube method, first presented by McKay et. al. [74]. The method works by ensuring that the intervals as projected onto each dimension of the design space are of equal length. Individual intervals along each dimension are then randomly grouped with intervals from other dimensions to form a single sample point. Due to the stochastic nature of this process, multiple runs of this method will yield different sample patterns for the domain. Furthermore, the space-filling of a particular pattern is not guaranteed so for a given sample point total it is necessary to compare the different Latin Hypercube designs to determine and select a design that achieves the best coverage of the domain. A commonly used method to improve the likelihood of a good space-filling pattern is to maximize the minimum distance between sample points as proposed by Morris et. al. [75]. This used the ‘maximin algorithm’ developed by Johnson et. al. [76] and aimed to find designs that offer a good compromise between the maximin criterion and the projective properties of the Latin Hypercube. Other methods have also been demonstrated, including the Audze-Eglais design [77, 78] which minimizes the “potential energy” of sampling points and a method by Park [79] which aims to maximize the sample point “entropy”.

In single-stage methods, such as Latin Hypercube the entire sample set is chosen before generating the data and creating the surrogate model. Adaptive sampling offers a potential improvement to customary single-stage or sequential space-filling methods by using information from the model.
itself to select successive new sample points. In strategies that involve adaptive sampling, after a small initial sample is made new sample points are added based on a set of criteria by analysing the samples collected so far. In this way, the sampling method can react to the shape of the function to achieve a better approximation to the data than single-stage space-filling methods.

In general adaptive strategies fall into one of two flavours, statistical and multi-criteria. The simplest form of statistical adaptive sampling involves estimating the error across the domain and adding new sample points at locations where the error is high. A variant of this method is used by Shewry et al. using a Kriging mean square error function to approximate the error and find potential sites for new samples.

By contrast, Multi-criteria approaches, as suggested by the name, use multiple criteria to determine the location of a new potential sample point. Often this involves just two criteria, one to promote space-filling and one to promote local refinement. An example of such a method is one developed by Jin et. al. based on the Leave One Out Cross-Point Validation (LOOCPV) error algorithm. The LOOCPV error is calculated throughout the domain and it is scaled based on the proximity to the current sample sites using a separation function. This provides a good balance between placing samples in regions of high perceived error as well as ensuring space-filling of the domain. However, to determine the LOOCPV error at potential sampling sites requires the introduction of interpolation models as demonstrated by Aute et. al. which adds an additional level of complexity.

Another example of a multi-criteria method is that developed by Tang et. al. In this method, the domain is divided into 'bins', with new sample points placed where the values of the model’s Laplacian are highest in each bin. This ensures a good spread of sample points across the domain while also conforming to the shape of the original function. However, this still results in excessive points in regions of low functional detail while insufficient points are placed where the functional detail is high or in areas of potential non-linearity. The method developed by Mackman et. al. attempts to address this by placing new sampling points based on both the Laplacian of the surrogate model and the separation function. This ensures that new points are placed in regions of high detail, while also ensuring a degree of space-filling to minimize uncertainty in less detailed regions. The advantage of this method is that the Laplacian can be calculated using the coefficients for the main model itself, thereby reducing the number of additional coefficients that must be calculated.

The adaptive sampling technique employed as part of this research was based on the work by Mackman et. al. This is a robust method designed specifically for surrogate models that use RBF interpolation. As a result of this, the process can be efficiently implemented as the method reuses matrices and coefficients from the RBF. This chapter will outline its construction and its relationship to the RBF interpolation. To verify its effectiveness the method will be compared to the single-stage Fully Factorial and Latin Hypercube Sampling methods.
3.3 Components of a Surrogate Model

There are three aspects to a surrogate model: the interpolation method, the sampling strategy, and the data source (otherwise known as the ‘black box’ function). As highlighted in Figure 3.1, these aspects are tightly linked: the data source is sampled by the sampling strategy, which in turn is interrogated by the interpolator. The results of the interpolation can then be compared to the original data source.

This section will outline the construction of the RBF interpolation method and the adaptive sampling strategy used and tested as part of this research. It will also outline the ‘black box’ functions used as the data sources during the development and testing of the surrogate model generator as well as methods used to determine the uncertainty of the resulting models.

Figure 3.1: Interactions of the components of a surrogate model
3.3.1 Interpolation

3.3.1.1 The Radial Basis Function Approach

The interpolation approach adopted for this body of work was the Radial Basis Function (RBF) \[64\]. The following section presents the general formulation of the RBF model. The code was implemented by the author, using the theory presented by Wendland \[84\], Fasshauer \[85\] and Mackman \[69\].

If \( y(x) \) is the original function to be approximated, let \( y_i \) be the result of the \( i \)th sample point \( x_i \), where \( x_i \) is a vector of the \( n \) dimensional design space such that \( x_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,n}) \). If the training data is taken to be a set of \( N \) samples, such that \( XX = x_1, x_2, \ldots, x_N \), with corresponding results \( YY = y_1, y_2, \ldots, y_N \), evaluating the RBF interpolation for the sample set \([XX, YY]\) at an arbitrary point \( x_* \) takes the form

\[
\hat{y}(x_*) = \sum_{i=1}^{N} \beta_i \phi(||x_* - x_i||) + p(x_*)
\]

(3.1)

where \( \beta_i, i = 1, \ldots, N \) are the model coefficients for the basis function \( \phi \), and \(||.||\) is used to denote the Euclidean norm. \( p(x) \) is an optional polynomial with \( M \) components with the general form

\[
p(x_*) = \sum_{k=1}^{M} \gamma_k f_k(x_*)
\]

(3.2)

where \( f_k(x) \) are the monomial components, and \( \gamma_k \) are the polynomial coefficients.

This provides \( N \) basis coefficient \( (\beta) \) and \( M \) polynomial coefficients \( (\gamma) \) that must be solved for using the training data \([XX, YY]\). This can be achieved by requiring the exact recovery of the data at each sample point such that \( \hat{y}(x_i) = y(x_i) \) for all data points \( x_i \) in \( XX \). An additional constraint, given by Equation \(3.3\) ensures that the polynomial is orthogonal to the basis functions.

\[
\sum_{i=1}^{N} \beta_i f_k(x_i) = 0, \quad k = 1, \ldots, M
\]

(3.3)

The problem can then be written in matrix form as follows

\[
\begin{bmatrix}
y \\
0
\end{bmatrix} =
\begin{bmatrix}
R & F \\
F^T & 0
\end{bmatrix}
\begin{bmatrix}
\beta \\
\gamma
\end{bmatrix}
\]

(3.4)

where

\[
R =
\begin{bmatrix}
\phi_{1,1} & \phi_{1,2} & \ldots & \phi_{1,N} \\
\phi_{2,1} & \phi_{2,2} & \ldots & \phi_{2,N} \\
\vdots & \vdots & \ddots & \vdots \\
\phi_{N,1} & \phi_{N,2} & \ldots & \phi_{N,N}
\end{bmatrix},
F =
\begin{bmatrix}
f_{1,1} & \ldots & f_{1,M} \\
f_{2,1} & \ldots & f_{2,M} \\
\vdots & \ddots & \vdots \\
f_{N,1} & \ldots & f_{N,M}
\end{bmatrix}
\]

(3.5)
3.3. COMPONENTS OF A SURROGATE MODEL

$R$ contains all pairwise combinations for the basis function such that $\phi_{i,j} = \phi(||x_i - x_j||)$ where $x_i$ and $x_j$ are both points in the training data $X$. $F$ contains the monomial results for each sample point $x_i$ such that $f_{i,k} = f_k(x_i)$. $\beta$ and $\gamma$ are the column vectors of coefficients for the basis function and polynomial respectively.

Obtaining the solution to this equation can cause issues with the conditioning of the system matrix, as a result of the large zero block. As described in Appendix B.1 the problem can be rearranged to give

$$\gamma = (F^TR^{-1}F)^{-1}F^TR^{-1}y$$

$$\beta = R^{-1}(y - F\gamma) = (R^{-1} - R^{-1}F(F^TR^{-1}F)^{-1}F^TR^{-1})y$$

Thus, for an arbitrary point $x_\ast$, the predicted result $\hat{y}$ can be determined as follows

$$\hat{y}(x_\ast) = r(x_\ast)^T\beta + f(x_\ast)^T\gamma$$

where $r$ and $f$ are vectors of the basis function and monomial terms for the point $x_\ast$, respectively, given by

$$r = \begin{bmatrix} \phi_{\ast,1} \\ \phi_{\ast,2} \\ \vdots \\ \phi_{\ast,N} \end{bmatrix}, \quad f = \begin{bmatrix} f_{\ast,1} \\ f_{\ast,2} \\ \vdots \\ f_{\ast,N} \end{bmatrix}$$

where $\phi_{\ast,i} = \phi(||x_\ast - x_i||)$ and $f_{\ast,k} = f_k(x_\ast)$.

3.3.1.2 Choice of basis function $\phi$

The choice of basis function ($\phi$) was critical to the effectiveness of this method. Many examples of potential basis functions exist, including linear and polynomial, as well as more complex functions such as Gaussian and multi-quadratic. The latter have additional degrees of freedom that allow for more precise fine-tuning of the function.

A further example of commonly used basis functions are the Wendland's functions [84]. For a stated number of continuous derivatives $C^2_k$ in $n$ dimensions, the Wendland’s functions are compact functions of minimal degree which decay to zero at a given distance from the centre. This distance is known as the support radius $R_s$, and allows the model to be tuned to achieve better accuracy. For this body of work Wendland’s $C^2$ functions were used as they provide a compromise between system matrix conditioning$^1$ and modelling behaviour [86]. Equation 3.17 gives the general form of

---

$^1$The conditioning of a system matrix measures the sensitivity of the output values to small changes in the input value. The conditioning number of a matrix can be considered to be the gain on errors in the input. This is explored in more detail in Section 3.4.2.
Table 3.1: Common examples of radial basis functions

<table>
<thead>
<tr>
<th>Basis Function</th>
<th>$\phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>linear</td>
<td>$1 - r$</td>
</tr>
<tr>
<td>Gaussian</td>
<td>$\exp(-r^2)$</td>
</tr>
<tr>
<td>multiquadric</td>
<td>$\sqrt{1 + r^2}$</td>
</tr>
<tr>
<td>inverse quadratic</td>
<td>$\frac{1}{1 + r^2}$</td>
</tr>
<tr>
<td>inverse multiquadric</td>
<td>$\frac{1}{\sqrt{1 + r^2}}$</td>
</tr>
<tr>
<td>$C^0$</td>
<td>$\begin{cases} \frac{1}{(l + 1)(l + 1)}, &amp; 0 \leq r \leq 1 \ 0, &amp; r &gt; 1 \end{cases}$</td>
</tr>
<tr>
<td>$C^2$</td>
<td>$\begin{cases} (1 - r)^l (l + 1), &amp; 0 \leq r \leq 1 \ 0, &amp; r &gt; 1 \end{cases}$</td>
</tr>
</tbody>
</table>

Wendland’s $C^2$ for $n$ spatial dimensions, where $l = \lfloor n/2 \rfloor + 2$.

$$\phi_{n,1} = \begin{cases} (1 - r)^l + 1 (l + 1)(l + 1), & 0 \leq r \leq 1 \\ 0, & r > 1 \end{cases}$$ (3.17)

The scaled distance from the sample point ($r$) is given by the Euclidean norm between the evaluation point $x_*$ and the control point $x_i$, and scaled by the support radius $R_s$. For the Wendland’s functions, the support radius indicates the limit of the influence of the function such that at the support radius the value of $\phi_{n,1} = 0$ as given by Equation (3.18). For the other functions, such as the Gaussian, the support radius is more akin to a standard deviation so while they influence all sample points, by $3R_s$ the influence is 0.44%.

$$r = \frac{||x_* - x_i||}{R_s}$$ (3.18)

### 3.3.2 Sampling Strategy

#### 3.3.2.1 Introduction

An important aspect of the surrogate model generation is the selection of the sampling points within the design space that will make up the data set to teach the model. The most notable methods include [Optimized Latin Hypercube Sampling (OLHS)] and orthogonal array (also referred to as [Fully Factorial (FF) sampling)] as described by Simpson et al. [67]. These methods are single-stage whereby all sample points are selected before the simulations/data gathering is performed. When correctly implemented, these methods can provide consistent coverage of the whole domain, though may require a relatively large number of samples to achieve a low overall error.
3.3. COMPONENTS OF A SURROGATE MODEL

(a) Fully Factorial Sampling
(b) Optimized Latin Hypercube Sampling
(c) Optimized Latin Hypercube Sampling + Adaptive Sampling

Figure 3.2: Example sampling strategies

One alternative is to implement a multi-stage adaptive sampling process. The goal of these adaptive sampling methods is to achieve a better distribution of sampling points to more accurately and efficiently capture the shape of the design space. These processes begin by taking an initial sample of the domain using one of the methods above and collecting the data for these sample points, then generating new sample points using some criteria based on the previously gathered data [60]. This process is repeated until the maximum allocation of sample points is reached. A comparison of the three schemes, Fully Factorial (FF), Optimized Latin Hypercube Sampling (OLHS), and Adaptive Sampling (with OLHS initial sample) are provided in Figures 3.2a, 3.2b & 3.2c respectively.

3.3.2.2 Fully Factorial Sampling

The simplest single-stage method is to use an orthogonal grid of sampling points sometimes referred to as Fully Factorial (FF) sampling [67]. This method has been demonstrated to perform well at returning low maximum error values, especially as it enforces the placement of samples at the corners of the domain and along its boundary.

3.3.2.3 Optimized Latin Hypercube Sampling

A popular method for achieving space-filling is the Latin Hypercube method. This works by ensuring that the samples, as projected onto each dimension of the design space, are equally spaced along them. Samples along each dimension are then randomly grouped with samples from other dimensions to form the sample points. On its own this does not guarantee good coverage of the domain space, so for a given number of points it is necessary to compare the different Latin Hypercube designs to determine and select the degree of space-filling. A commonly used method to improve the likelihood of space-filling, and the one used here, is to maximize the minimum distance between sample points as proposed by Morris et. al. [75].

Unlike a Fully Factorial sampling scheme, the Optimized Latin Hypercube Sampling (OLHS) method alone does not guarantee good coverage of the sample space up to the boundary. Therefore, to ensure the surrogate model is valid up to the edge of the design space, the initial sample set must include samples from key locations along it. The common practice is to use central composite design which allocates samples to the corners of the domain space as well as at the centre of each edge, face and cell. By combining OLHS with ‘central composite design’, the hybrid sampling
scheme can effectively fill the design space with samples while also ensuring good coverage of the domain boundary.

The samples in this hybrid approach are divided between those that are allocated using the central composite design and those that are allocated using OLHS. This works well for problems with few dimensions, but the number of samples allocated to the boundary increases with $O(3^n)$ for the number of parameters $n$ as can be seen in Table 3.2. For adaptive sampling methods, such as those described in Section 3.3.2.4, this can mean significantly fewer samples are allocated to the initial space-filling scheme, so careful consideration of the initial and overall sample size is required. For the sake of clarity, this combined method will continue to be referred to as Optimized Latin Hypercube Sampling (OLHS) throughout this work.

Table 3.2: Design space boundary sample allocation and source per number of parameters

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertices</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>16</td>
<td>32</td>
<td>64</td>
</tr>
<tr>
<td>Edges</td>
<td>1</td>
<td>4</td>
<td>12</td>
<td>32</td>
<td>80</td>
<td>192</td>
</tr>
<tr>
<td>Face</td>
<td>-</td>
<td>1</td>
<td>6</td>
<td>24</td>
<td>80</td>
<td>240</td>
</tr>
<tr>
<td>Cell</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>8</td>
<td>40</td>
<td>160</td>
</tr>
<tr>
<td>4-face</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>10</td>
<td>60</td>
</tr>
<tr>
<td>5-face</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>6-face</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>9</td>
<td>27</td>
<td>81</td>
<td>243</td>
<td>729</td>
</tr>
<tr>
<td>V+E+F</td>
<td>3</td>
<td>9</td>
<td>26</td>
<td>72</td>
<td>192</td>
<td>496</td>
</tr>
</tbody>
</table>

3.3.2.4 Adaptive sampling

The goal of adaptive sampling is to achieve a better distribution of sampling points to more accurately and efficiently capture the shape of the design space. These processes begin by taking an initial sample of the domain, using one of the single-stage methods outlined above, and collecting the data for those sample points. New sample points can then be generated using some criteria based on the previously gathered data [60]. This process is repeated until the maximum allocation of sample points is reached. According to Forrester et. al. [58] one-third of the total sampling points, or as close as the method employed would allow, should be allocated to the initial sample.

The adaptive strategy employed for this research used a multi-criteria method based on the work of Mackman et. al. [69]. The criterion they devised is the product of the Laplacian and the separation function and thus gives a balance between adding points in locations of high detail and unsampled regions of the domain, thus ensuring a degree of space-filling to minimize uncertainty in less detailed regions. Their criterion $C$ is given by

$$C(x^*) = (|\nabla^2 \hat{y}(x^*)| + \epsilon)(1 - h(x^*))^2$$  (3.19)
where $\nabla^2 \hat{y}$ is the Laplacian of the surrogate model $\hat{y}$, $c$ is an offset parameter to ensure a non-zero value when $|\nabla^2 \hat{y}(\mathbf{x}_*)| = 0$, and $h(\mathbf{x})$ is the separation function defined such that $(1 - h)$ grows with increased distance from the data sites. Large values of $C$ indicate potential new sampling sites.

The Laplacian of the interpolation model is defined using Equation 3.20. Applying this to Equation 3.1, the derivative of the model for a given parameter $x_i$ can be evaluated using Equation 3.21. For a given data set, the coefficients $\beta$ and $\gamma$ remain constant, thus evaluations of the Laplacian of the data is no more expensive than evaluations of the model itself.

$$\nabla^2 \hat{y} = \sum_{j=1}^{N} \frac{\partial^2 \hat{y}}{\partial x_j^2}$$ (3.20)

$$\frac{\partial^2 \hat{y}}{\partial x_j^2} = \frac{\partial^2 \mathbf{r}^T}{\partial x_j^2} \beta + \frac{\partial^2 \mathbf{f}^T}{\partial x_j^2} \gamma$$ (3.21)

The separation function $h$ can be constructed as an RBF surrogate model allowing a similar framework to be employed to that of the primary surrogate model. In this instance, the data sites are given a value of 1 with the separation function decaying with increased distance from these points. Wendland’s $C^2$ (Equation 3.17) was chosen for the radial function based on Mackman et. al. [69] who suggested that a smooth function such as the Wendland’s function would provide better sampling as they would blend better with the Laplacian term. Given that Wendland’s $C^2$ decays to zero at the support radius, the separation function $h$ lies in the range $0 \leq h \leq 1$ everywhere in the domain, with a value of one indicating a point coincident with a sample location and a value of zero denoting a region far from all sample points.

As with the primary surrogate model, it is important to set an appropriate support radius for the separation function, though in this case, the $h$ should decay to a minimum (ideally zero) at a point that is furthest from all points on the domain. This can be achieved by setting the support radius proportional to the fill distance. The fill distance is the largest nearest neighbour distance of the current set of samples, given by

$$d = \max\{i = [1, 2, ..., N] \mid \min \|\mathbf{x} - \mathbf{x}_i\|\}$$ (3.22)

and the support radius used in the separation function interpolation is then some scalar constant ($k$) times the fill distance

$$R = kd$$ (3.23)

Each time a new sample point is added the support radius of the separation function is updated.

Figure 3.3 shows an example of a new sample point being added to a surrogate model of a one-dimensional slice of the droplet function (Equation 3.24) using this method. Figure 3.3a provides a comparison between the original function and the surrogate model with the current set of sample points. The error at this stage is shown in Figure 3.3b.
CHAPTER 3. GENERATION AND VALIDATION OF A SURROGATE MODEL FOR RAREFIED GAS MODELLING

Figure 3.3: Adding a new sample to the surrogate model using the adaptive criterion in Equation 3.19
As described above, the criteria for identifying new sites include two parts, the separation function and the Laplacian. The separation function \((1 - h(\mathbf{x}))^2\) for the current number of samples is provided in Figure 3.3c. From this figure, the remotest areas of the domain are currently between samples -1 and -0.34. The Laplacian of the surrogate model with the current samples is shown in Figure 3.3d. As might be expected for radial functions, the curvature peaks at the sample sites with the highest curvature coincident with the central droplet. The criterion given by Equation 3.19 is shown in Figure 3.3e. Large values of \(C\) indicate potential new sampling sites, in this case, the peak value occurs at -0.6, so this is chosen as the new sample for the model.

Figure 3.3f shows the surrogate model with the new sample included. As can be seen from this figure and Figure 3.3g, this new surrogate model has a better agreement with the original function. It should be noted that absolute agreement between the surrogate model and original function is still low, however, this is due to the low number of samples which was kept low for the sake of clarity.

3.3.3 The Function

3.3.3.1 Test cases

To be effective, a surrogate model should reproduce the original ‘black box’ function as accurately as possible. To verify this it is crucial to test and validate the model against a series of known, stable functions that are easy to compute but provide a challenge to the generation process. A function that is easy to compute is essential during the development cycle as it allows multiple rounds of rapid testing as bugs in the code or weaknesses in the algorithms are identified and resolved. It also allows the fast generation of a high fidelity evaluation grid against which the surrogate model can ultimately be judged.

Four test cases were selected to facilitate this verification process, two analytical functions and two aerodynamic test cases using panel methods. All four test cases have two input parameters to enable easier visual inspection of the generated surfaces such that abnormalities may be spotted which may not be apparent in the numerical metrics.

The two analytical functions that were used are the Droplet function \(D(x, y)\) and Franke’s function \(F(x, y)\) given by Equations 3.24 & 3.25 respectively. The expected surfaces for the two functions are shown in Figures 3.4a & 3.4b respectively. These functions were used by Mackman et al as examples of challenging test cases to help examine different adaptive sampling regimes. They have been reused here for the same reason and to provide a point of comparison with their work. The droplet function for instances has regions of high curvature near the centre of the domain and in a ring encircling the centre. These details could be difficult for the surrogate model to sample effectively and lead to large errors in these regions of the domain. By contrast, Franke’s function has a lower maximum curvature but does have multiple peaks with significant changes in curvature across the domain. This will therefore test the balance of sampling in regions with ‘higher’ detail and ensuring the domain as a whole is effectively covered.
CHAPTER 3. GENERATION AND VALIDATION OF A SURROGATE MODEL FOR RAREFIED GAS MODELLING

\[ D(x, y) = -4e^{-\frac{25}{8}(x^2+y^2)} + 7e^{-\frac{125}{4}(x^2+y^2)} \]  

\[ F(x, y) = \frac{3}{4}e^{-\frac{1}{4}(9x-2)^2+(9y-2)^2} + \frac{3}{4}e^{-\frac{1}{10}(9x+1)^2-\frac{1}{10}(9y+1)^2} + \frac{1}{2}e^{-\frac{1}{4}((9x-7)^2+(9y-3)^2)} - \frac{1}{5}e^{-(9x-4)^2-(9y-7)^2} \]  

The second two test cases are based on simulated data for a simple two-dimensional body in a rarefied gas with aspects of its geometry being varied. Test case AERO-1 assumes a simple rectangular satellite body whose length and height are varied. In the case of AERO-2, a fixed rectangular body with a height of 0.5m and a length of 1m is assumed while the geometry of the nose is varied. In this instance, a simple conic nose is used and can be defined by its length and radius at the front tip (see Figure 3.5). For each of the aerodynamic test cases the \( C_D \) was calculated using Sentman’s analytical equations for rarefied air (see Section 2.4.3.5 for further details). While less accurate than DSMC methods, these panel methods provide a fast and repeatable result and are therefore ideal for testing the validity of the surrogate model. The ideal surface for test cases AERO1 and AERO2 are shown in Figures 3.4c & 3.4d respectively.

(a) Droplet function

(b) Franke’s function

(c) AERO-1: Varying length and height

(d) AERO-2: Fixed dimensions and varying nose geometry

Figure 3.4: Reference data for the four test cases
3.3. COMPONENTS OF A SURROGATE MODEL

3.3.4 Measuring Uncertainty

Measuring the uncertainty of the model is an important aspect of validating the method while also ensuring the generated models are sufficiently accurate. What follows is an outline of three methods that were used to measure the uncertainty; Gridded Evaluation, [LOOCPV] and Power Density Uncertainty. For all three methods, the level of uncertainty in the generated model is calculated using the Root Mean Squared Error (RMSE) given by:

\[ RMSE = \sqrt{\frac{1}{P} \sum_{i=1}^{P} (\hat{y}(\mathbf{x}_i) - y(\mathbf{x}_i))^2} \]  \hspace{1cm} (3.26)

where \( \hat{y} \) and \( y \) are the output of the model and the true value of the 'black box' function respectively at point \( i \) out of \( P \) evaluation points. The methods differ primarily in the selection of the evaluation points and the method by which the error \( (\hat{y}(\mathbf{x}_i) - y(\mathbf{x}_i)) \) is calculated.

3.3.4.1 Gridded Evaluation

Perhaps the simplest and most representative uncertainty measure is to compare the generated surrogate model against a set of pre-calculated results. Typically, the evaluation points for these pre-calculated results are arranged in a gridded pattern across the whole domain thus providing consistent and even coverage. While other space-filling methods can also be used, the key element is that this set is known and consistent for a given 'black box' function. This means the error values are directly comparable between surrogate model generation methods.

The relationship between the number of evaluation points \( P \) and the RMSE has been shown to be broadly asymptotic[69] and, more generally, consideration should be made for the compatibility with the methods being examined. For instance, this could mean selecting an evaluation grid whose samples per dimension were a large prime number, to minimize the number of evaluation points coincident with samples (especially if the gridded sample is used) and resonance with the function. However, a \( 51 \times 51 \) evaluation grid was chosen as this allowed comparison with previous work in this field and thus aided in bug tracking in the code. This does mean that \( 3^2 \) and \( 13^2 \) evaluation grids will have a degree of resonance within the evaluation grid and thus the observed uncertainty may be less accurate.

Since this method requires two distinct sets of samples, one to train the surrogate model and one to test the model, the method is not very efficient for determining the uncertainty of a surrogate model of an unknown 'black box' function. This problem can be minimized by limiting the size...
of the evaluation set. However, these evaluation samples may not be used in the model despite being collected. Other methods that do not rely on a separate evaluation set such as LOOCPV and the Power Function may be more appropriate in this case. The Gridded Evaluation method was therefore used exclusively for testing and validating different methods of surrogate model generation.

3.3.4.2 Leave One Out Cross-Point Validation (LOOCPV)

When determining the uncertainty of a surrogate model, it is useful to compare the results of the model to the original ‘black box’ function. However, as was seen with the Gridded Evaluation this would normally require two sets of samples: one to train the surrogate model and one to evaluate it. This is not particularly efficient especially if the evaluation sample points are not used in the model itself. Leave One Out Cross-Point Validation (LOOCPV) addresses this by allowing a single data set to be used for both the training and validation of the surrogate model.

Under LOOCPV a number of evaluation points were selected from the training data set (selection criteria is discussed below). Each evaluation point was then excluded from the training data in turn and the matrices and coefficients of the reduced surrogate model were recalculated. This reduced surrogate model was then interrogated at the location of that evaluation point and the result was compared to its known value. The RMSE could then be calculated for the complete set of evaluation points.

It is important to consider which samples from the training data set should be considered as evaluation points. For instance, samples that lie along the domain boundary should not be used as evaluation points since these boundary sample points ensure that the model is valid right up to the edges of the design space, as is discussed in Section 3.3.2.3.

The key drawback with LOOCPV is the computational cost of calculating the error, which for \( N \) samples is \( O(N^4) \). This can be tempered by having a smaller fixed number of evaluation points but this simply trades off the accuracy of the uncertainty value with computational time. Similarly, an interpolation model of the LOOCPV error can be used and has been shown to achieve \( O(N^3) \) but this introduces an additional level of approximation.

3.3.4.3 Power Function

A common method for estimating the uncertainty of an RBF interpolation scheme can be constructed based on the so-called power function and a factor depending on the interpolant itself. The power function \( P_{\phi,\chi}(x_*) \) (given by Equation 3.27) quantifies the component of interpolation error that is dependent on the basis function, model parameters and sample locations.

\[
P_{\phi,\chi}(x_*) = \sqrt{\phi(0) + u^T(F^T R^{-1} F)^{-1} u} - r^T R^{-1} r
\]

(3.27)

\[
u = F^T R^{-1} r - f
\]

(3.28)
Jakobsson et. al. showed that an estimation of the local residual error can be made using Equation 3.29 where \( m \) is an interpolant dependent multiplier and \( \alpha \) is an exponent. In their paper, Jakobsson et. al. tested a number of multipliers candidates but the most appropriate for the work being carried out here is the inverse of the absolute value of the interpolant \( \frac{1}{|\hat{y}|} \). 

\[
    r(x_*) = P_{\Phi,\lambda}(x_*)m(x_*)^\alpha
\]

(3.29)

In general \( r(x) \geq |\hat{y}(x) - y(x)| \) thus the residual can be used with Equation 3.26 to find the maximum RMSE for the surrogate model. To provide a consistent and even coverage of the domain the evaluation points are arranged in a gridded pattern except for the boundary. To avoid being coincident with a sample point at the boundary (which would return a residual of 0), points are not evaluated along the boundary. This means that the method has a computational complexity similar to the Gridded Evaluation of \( O(N_E^3) \) where \( N_E \) is the number of evaluation points. Therefore of the two uncertainty measures that do not require a dedicated testing data set, the Power Function scheme tends to be much faster than LOOCPV.
3.4 Initial Observations and Method Refinement

3.4.1 Choosing a Basis Function $\phi$

As described in Section 3.3.1.2, an important aspect of the RBF method is the choice of basis function. Table 3.1 illustrates that there are many examples of basis functions, including linear and polynomial, as well as more complex functions such as Gaussian and multi-quadratic. The latter have additional degrees of freedom that allow for more precise fine-tuning of the function. This section will examine the performance of these basis functions and assess which one will be most appropriate for the work performed here.

Using the four test cases described in Section 3.3.3.1 (Droplet Function, Franke Function, Aero-1 and Aero-2), a series of surrogate models were generated for each of the basis functions listed in Table 3.1. These functions were linear (Equation 3.10), Gaussian (Equation 3.11), multiquadric (Equation 3.12), inverse quadric (Equation 3.13), inverse multiquadric (Equation 3.14) as well as

![Figure 3.6: Comparison of RMSE for each basis function with varying sampling budgets](image_url)
Wendland’s $C^0$ & $C^2$ functions for two dimensions (Equations 3.15 & 3.16 respectively). In each case, the Fully Factorial (FF) sampling strategy was used (as described in Section 3.3.2) with a sample budget of between $5^2$ and $13^2$. To calculate the RMSE, the resulting surrogate models were compared against a known 51x51 grid of each test case. For the Wendland’s functions, an $R_s$ of 1 was used, while for the other functions an equivalent $R_s = \frac{1}{3}$, with the difference in value arising from what $R_s$ represents for each basis function. For the Wendland’s functions, the support radius indicates the limit of the influence of the function, however, for the other functions (such as the Gaussian) the support radius is more akin to the standard deviation.

Figure 3.6 shows the RMSE of the basis function for the four test cases. In general, it can be seen from the figure that in all cases the uncertainty tends to improve as the number of samples is increased. Wendland’s $C^2$ performs well across all four test cases, though it is observed that the multiquadric, inverse quadric and inverse multiquadric achieved a lower uncertainty for the

![Figure 3.6: Comparison of reciprocal condition number of the system matrix for each basis function with varying sampling budgets](image)

Figure 3.7: Comparison of reciprocal condition number of the system matrix for each basis function with varying sampling budgets
Aero test cases. The Gaussian by comparison was relatively unreliable. While it performed well compared to the other basis functions for the Droplet and Aero-1 test cases, it struggled to reduce the uncertainty of Franke’s function. Similarly, while initially achieving the lowest uncertainty in the Aero-2 test case, the Gaussian basis function became unstable as more samples were added.

The instability in the Gaussian result is likely due to the conditioning of the system matrix \( R \). Due to their formulation, some of the basis functions have a global influence over the domain, even when the \( R_{i} \) is small. While this does not necessarily imply that a single point will have a strong influence over all other points, their influence is still non-zero which impacts the conditioning of \( R \). Plotting the reciprocal conditioning number, as is seen in Figure 3.7, gives a sense of the smallest perturbation which will affect the most significant figure of the result. In this instance, a value of 1 indicates a well-conditioned matrix, while a value of 0 indicates that the matrix is singular and not invertible. Thus, a matrix that has a small reciprocal condition number is not well-conditioned.

As can be seen from Figure 3.7, the Gaussian basis function is generally very poorly conditioned which explains its unreliable performance at reproducing the original test cases. Similarly, the reciprocal conditioning numbers of the multiquadric, inverse quadric, and inverse multiquadric are also very low. Since the intent is to sample DSMC simulations that can have a random component to their results, it is prudent to maximize the reciprocal conditioning number where possible. So, while the Wendland’s functions did not produce the lowest uncertainties, they did perform consistently well across all 4 test cases and maintained a high reciprocal conditioning number in the process. For this reason, the Wendland’s \( C^2 \) function was chosen for this work.

### 3.4.2 Choosing a Support Radius

As described in the previous section, the basis function chosen for the work performed as part of this thesis was the Wendland’s \( C^2 \) function. An example of a Wendland’s \( C^2 \) function for 2 dimensions is given in Equation 3.17. These functions decay to zero at a given distance from the centre known as the support radius \( R_{s} \). The support radius controls the area of the domain over which a given sample point has influence.

Figure 3.8 outlines the key trade-off when selecting a support radius. If the radius is too small, then the interpolated surface may ‘sag’ in the regions between sample points. However, as the support radius gets larger, a sample point may achieve global coverage of the domain and thus have an influence on all other sample points. If the support radius gets sufficiently large that all points have a strong global influence, then the system matrix \( R \) (Equation 3.5), becomes poorly conditioned and accurate recovery of the model coefficients becomes difficult. The radius at which a sample obtains global coverage depends on its location. However, all samples will have global influence when the support radius is equal to the length of the longest diagonal of the normalized sample space hypercube, which for two dimensions/parameters is \( \approx 1.41 \).

A series of surrogate models were generated with a normalized support radius between 0.1 and 2 using the four test cases described in Section 3.3.3.1 (Droplet Function, Franke Function, Aero-1 and Aero-2). For each test case, the four sampling strategies described in Section 3.3.2 were used (Fully Factorial (FF), Optimized Latin Hypercube Sampling (OLHS) and Fully Factorial (FF) +
Support Radius

Small

Surrogate Model

Large

System Matrix Conditioning

Well Conditioned

Poorly Conditioned

Figure 3.8: Trade-off between reproducing the function and the conditioning of the systems matrix as the support radius changes. (Note: the size of the bubbles in the matrix indicate the amount of influence one point has over the other)

Adaption and [OLHS] + Adaption, to identify if there was any dependency on the method when selecting the support radius. The generation process required a total of 100 samples each and the resulting models were compared against a 51x51 grid for each test case. The RMSE for each scenario is plotted against the support radii in Figure 3.9.

For the majority of the scenarios in Figure 3.9, the RMSE appears to improve until the $R_s$ reaches a value of 1 (the length of the domain) where the uncertainty measures approach a lower limit. Above an $R_s$ of 1, there is no substantial change or improvement in RMSE, thus suggesting that to minimize the uncertainty in the model, the $R_s$ should be at least 1.
CHAPTER 3. GENERATION AND VALIDATION OF A SURROGATE MODEL FOR RAREFIED GAS MODELLING

An $R_s$ greater than 1 implies that any given sample point will likely influence a large proportion of the other samples in the domain. While this does not necessarily imply a strong influence over all of these points, their influence is still non zero which impacts the conditioning of the system matrix $R$. Plotting the reciprocal conditioning number, as is seen in Figure 3.10, gives a sense of the smallest perturbation which will affect the most significant figure of the result. In this instance, a value of 1 indicates a well-conditioned matrix, while a value of 0 indicates that the matrix is singular and not invertible. Thus a matrix that has a small reciprocal condition number is not well-conditioned.

As can be seen from Figure 3.10 the system matrix $R$ is generally not well conditioned. However, for the radii tested here, the conditioning was more than sufficient to limit the impact of floating-point round-off errors (for more details see Appendix B.3). Given this, it was decided to use a support radius of 1.41 (the length of the diagonal) with the more general $R_s = \sqrt{n}$ for higher dimensions.

Figure 3.9: Comparison of RMSE for each sampling strategies with varying support radii and a fixed sample budget of 100.
3.4. INITIAL OBSERVATIONS AND METHOD REFINEMENT

Figure 3.10: Reciprocal conditioning number for each sampling strategies with varying support radii and a fixed sample budget of 100

As was seen from Figure 3.9, this ensures that the uncertainty values have attained their lower limits for the respective test cases and methods. Choosing the length of the diagonal rather than 1 also provided a degree of consistency by enforcing global influence on all samples. This in turn also ensured the good recovery of the surface even when there are few sample points, such as during the initial sampling stages of adaptive schemes.

3.4.3 Order of the Polynomial

In addition to the basis function, the RBF method has the option of including a polynomial as given in Equation 3.1. The rationale is that the polynomial can account for the trend of the data while the radial basis function provides a form of correction in areas that have a low agreement with that trend.

Table 3.3 shows the number of components of a polynomial with order 0 to 5 and the number of variables between 1 and 5. As can be seen from the table, the number of components increases quickly with both the number of variables and the order of the polynomial. It is therefore important to know not only whether including a polynomial would be appropriate for the surrogate model in this setting but also what the order of polynomial should be used.

Table 3.3: Number of monomial components required for a polynomial of order 0 to 5 for a design space with 1 to 5 variables.

<table>
<thead>
<tr>
<th>Number of variables</th>
<th>Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 1 2 3 4 5</td>
</tr>
<tr>
<td>1</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>2</td>
<td>1 3 6 10 15 21</td>
</tr>
<tr>
<td>3</td>
<td>1 4 10 20 35 56</td>
</tr>
<tr>
<td>4</td>
<td>1 5 15 35 70 126</td>
</tr>
<tr>
<td>5</td>
<td>1 6 21 56 126 252</td>
</tr>
</tbody>
</table>
CHAPTER 3. GENERATION AND VALIDATION OF A SURROGATE MODEL FOR RAREFIED GAS MODELLING

To examine whether a polynomial would be suitable in this situation a set of scenarios were created using the four sampling strategies described in Section 3.3.2, (FF, OLHS and FF+Adaptation and OLHS+Adaptation) and the four test cases described in Section 3.3.3.1, (Droplet Function, Franke Function, Aero-1 and Aero-2). In each case, a series of surrogate models was generated, first without a polynomial and then using a polynomial with orders of 0 to 10. The generation process required a total of 100 samples each and used a support radius of $R_S = 1.41$. The resulting models were compared against a 51x51 grid of each test case. The RMSE for each scenario is plotted against the order of the polynomial in Figure 3.11a while Figure 3.11b shows the run time.

As can be seen from Figure 3.11a, there is no significant improvement in the RMSE including or excluding the polynomial for any of the scenarios tested. However, as can be seen from Figure 3.11b there is an increase in the run time for all scenarios as the order of the polynomial is increased. This would be compounded further should the design space require more than two parameters. As including a polynomial provides no obvious improvement in the accuracy of the model, it would not be appropriate to include it in this setting. A more complete analysis of the sampling strategies is provided in Appendix B.4.

3.4.4 Comparison of Sampling Strategies

An important aspect of the surrogate model generation is the choice of sampling strategy. Section 3.3.2 outlined a number of common sampling strategies including simple single-stage sampling methods, such as the Fully Factorial (FF) and Optimized Latin Hypercube Sampling (OLHS) to more complex methods that adapt to the function being examined. In the review paper by Forrester et. al. [59], adaptive sampling schemes were identified as a preferred sampling strategy for aerodynamic research. This section will summarise the work performed to assess whether adaptive methods are appropriate for the work being done as part of this thesis. A more complete analysis of the sampling strategies is provided in Appendix B.2.
Four sampling methods were considered: the two single-stage sampling methods: FF (Section 3.3.2.2) and OLHS (Section 3.3.2.3) and the Adaptive strategy (Section 3.3.2.4) using either FF (FF+Adapt) or OLHS (OLHS+Adapt) as the first stage. Figure 3.12 shows how the RMSE of the generated surrogate model varies as the sample budget is increased. This is shown for the sampling methods and the four test cases described in Section 3.3.3.1. A detailed breakdown of the model for a sample budget of 100 points is in Table 3.4. The RMSE for the surrogate models in Figure 3.12 and Table 3.4 were calculated using the Gridded Evaluation method described in Section 3.3.4.1 with a $51 \times 51$ evaluation grid.

As can be seen from Figure 3.12, for a given sample budget, the adaptive sampling strategy generally achieved the lowest overall RMSE. This is as expected and therefore confirms the validity of the approach, especially for the two aerodynamic test cases: Aero-1 (Figure 3.12c) and Aero-2 (Figure 3.12d). For the adaptive methods, there was no significant difference between using either FF or OLHS as the initial stage.

It is important to highlight that the fluctuation in the FF sampling methods for the Droplet Function (Figure 3.12a) were mostly due to whether a sample happens to be coincident with the central peak for a given sample budget. In all other cases, the FF had a strong negative correlation with the sample budget. In general, OLHS also followed a similar trend to the FF, but fluctuated around the gridded sampling method, which is likely due to the stochastic nature of the OLHS method.

While the adaptive methods did achieve an improvement in the RMSE as compared to the single-stage results, they were also significantly slower at selecting samples for the surrogate model (see Appendix B.2 for more details). However, since the aim is to use DSMC methods on the satellite bodies to determine the aerodynamic forces, it is important to consider the total run time required to generate the surrogate model, i.e: the combined time of selecting and generating all samples.

Table 3.4: A comparison of sampling strategies with maximum sampling points of 100

<table>
<thead>
<tr>
<th>Test Case</th>
<th>Sampling</th>
<th>min value</th>
<th>max value</th>
<th>RMSE</th>
<th>Max Error</th>
<th>Run time [s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Droplet</td>
<td>FF</td>
<td>-2.637</td>
<td>0.60</td>
<td>$2.07 \times 10^{-1}$</td>
<td>$2.40 \times 10^{-1}$</td>
<td>2.17</td>
</tr>
<tr>
<td></td>
<td>OLHS</td>
<td>-2.746</td>
<td>3.00</td>
<td>$1.26 \times 10^{-1}$</td>
<td>$1.05 \times 10^{-1}$</td>
<td>2.31</td>
</tr>
<tr>
<td></td>
<td>FF + Adapt</td>
<td>-2.680</td>
<td>3.00</td>
<td>$2.15 \times 10^{-2}$</td>
<td>$1.48 \times 10^{-1}$</td>
<td>59.16</td>
</tr>
<tr>
<td></td>
<td>OLHS + Adapt</td>
<td>-2.683</td>
<td>3.00</td>
<td>$2.13 \times 10^{-2}$</td>
<td>$1.77 \times 10^{-1}$</td>
<td>59.69</td>
</tr>
<tr>
<td>Franke</td>
<td>FF</td>
<td>-0.197</td>
<td>1.037</td>
<td>$1.23 \times 10^{-3}$</td>
<td>$9.73 \times 10^{-3}$</td>
<td>2.05</td>
</tr>
<tr>
<td></td>
<td>OLHS</td>
<td>-0.187</td>
<td>1.039</td>
<td>$6.02 \times 10^{-3}$</td>
<td>$4.25 \times 10^{-2}$</td>
<td>2.12</td>
</tr>
<tr>
<td></td>
<td>FF + Adapt</td>
<td>-0.190</td>
<td>1.039</td>
<td>$3.56 \times 10^{-3}$</td>
<td>$2.06 \times 10^{-2}$</td>
<td>59.31</td>
</tr>
<tr>
<td></td>
<td>OLHS + Adapt</td>
<td>-0.188</td>
<td>1.039</td>
<td>$2.70 \times 10^{-3}$</td>
<td>$1.36 \times 10^{-2}$</td>
<td>59.26</td>
</tr>
<tr>
<td>Aero-1</td>
<td>FF</td>
<td>-8.485</td>
<td>-2.118</td>
<td>$2.60 \times 10^{-1}$</td>
<td>$2.02 \times 10^{0}$</td>
<td>7.10</td>
</tr>
<tr>
<td></td>
<td>OLHS</td>
<td>-8.485</td>
<td>-1.826</td>
<td>$3.13 \times 10^{-1}$</td>
<td>$3.05 \times 10^{0}$</td>
<td>6.67</td>
</tr>
<tr>
<td></td>
<td>FF + Adapt</td>
<td>-8.485</td>
<td>-2.107</td>
<td>$9.24 \times 10^{-2}$</td>
<td>$1.19 \times 10^{0}$</td>
<td>118.41</td>
</tr>
<tr>
<td></td>
<td>OLHS + Adapt</td>
<td>-8.485</td>
<td>-2.096</td>
<td>$1.14 \times 10^{-1}$</td>
<td>$1.04 \times 10^{0}$</td>
<td>110.83</td>
</tr>
<tr>
<td>Aero-2</td>
<td>FF</td>
<td>-2.384</td>
<td>-2.192</td>
<td>$1.86 \times 10^{-3}$</td>
<td>$1.17 \times 10^{-2}$</td>
<td>7.32</td>
</tr>
<tr>
<td></td>
<td>OLHS</td>
<td>-2.384</td>
<td>-2.192</td>
<td>$3.08 \times 10^{-3}$</td>
<td>$2.49 \times 10^{-2}$</td>
<td>7.27</td>
</tr>
<tr>
<td></td>
<td>FF + Adapt</td>
<td>-2.385</td>
<td>-2.192</td>
<td>$9.08 \times 10^{-4}$</td>
<td>$8.75 \times 10^{-3}$</td>
<td>115.28</td>
</tr>
<tr>
<td></td>
<td>OLHS + Adapt</td>
<td>-2.385</td>
<td>-2.192</td>
<td>$8.34 \times 10^{-4}$</td>
<td>$6.86 \times 10^{-3}$</td>
<td>116.77</td>
</tr>
</tbody>
</table>
CHAPTER 3. GENERATION AND VALIDATION OF A SURROGATE MODEL FOR RAREFIED GAS MODELLING

(a) Droplet Function

(b) Franke’s function

(c) AERO-1: Varying length and height

(d) AERO-2: Fixed dimensions and varying nose geometry

Figure 3.12: Comparison of RMSE for each sampling strategies with varying number of samples and fixed support radius $R_s = \sqrt{2}$

Figure 3.14 shows the RMSE plotted against the generation time assuming a DSMC simulation time of 1 hour per sample (see section Section 2.5.3) for the aerodynamic test cases. Parallel computation has not been accounted for in this case, but would result in the same improvement in run time and overheads for all methods. As can be seen from Figure 3.14, the adaptive methods are faster for a given uncertainty and more accurate for a given run time, assuming a DSMC simulation time. In general, it was observed that, so long as the time taken to determine a new sample point was much smaller than the time taken to collect the sample, the adaptive methods were always better for a given surrogate model uncertainty.
3.4. INITIAL OBSERVATIONS AND METHOD REFINEMENT

Figure 3.13: Generated surrogate model and absolute error in the model compared to the true function of the Droplet function for the 4 sampling strategies for $N = 100$
CHAPTER 3. GENERATION AND VALIDATION OF A SURROGATE MODEL FOR RAREFIED GAS MODELLING

In summary, for a given sample budget, the adaptive sampling method, as implemented in the code, achieves a better uncertainty measure than the single-stage space-filling methods in most cases. However, this comes with an additional performance penalty for the generation time when compared to other methods. For analytical methods, this means it is often faster to run a single-stage sampling method with a higher sample budget, than an adaptive scheme for a desired level of uncertainty. Nonetheless, if acquiring the sample takes significantly longer than it takes to identify new sample locations using the adaptive criteria, such as in the case of DSMC simulations, then adaptive methods are the faster approach for a given uncertainty measure. In general, this aligns with observations in work by other authors[86].

3.4.5 Sample Budget Easing to improve RMSE

The sampling strategies described in Section 3.3.2 and tested in the section above have relied on a predefined sample budget before the model is generated. This works well for comparing methods on known functions. However, for an unknown function, it can be hard to predict what an appropriate sample budget might be, specifically, a sample budget that achieves the desired level of uncertainty in the surrogate model for the work being performed.

An initial approach to selecting the sample budget could be to base the budget on previously generated surrogate models. However, as was seen in Figure 3.12, the surrogate models generated for each of the test cases had very different uncertainty measures for a given sample budget. By extension, if a certain level of uncertainty was required, the sample budget to achieve this would vary significantly depending on the 'black box' function. While estimates could be made, from initial work with actual data, it became apparent that a means of relaxing the budget was necessary.
Table 3.5: \( R^2 \) Goodness-of-fit of a number curves to the RMSE values in \( \log_{10} \) space for each test case

<table>
<thead>
<tr>
<th>Models</th>
<th>Droplet</th>
<th>Franke</th>
<th>Aero-1</th>
<th>Aero-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>linear ax + b</td>
<td>0.9040</td>
<td>0.9029</td>
<td>0.9772</td>
<td>0.5238</td>
</tr>
<tr>
<td>exponential ( a \exp(bx) )</td>
<td>0.8421</td>
<td>0.8622</td>
<td>0.9675</td>
<td>0.5147</td>
</tr>
<tr>
<td>power law ( ax^b )</td>
<td>0.9756</td>
<td>0.9894</td>
<td>0.9509</td>
<td>0.7784</td>
</tr>
<tr>
<td>power law + constant ( ax^b + c )</td>
<td>0.9900</td>
<td>0.9955</td>
<td>0.9775</td>
<td>0.9748</td>
</tr>
</tbody>
</table>

A method of dynamically easing the sample budget was implemented to help improve the uncertainty in the model. Firstly, a ‘guess’ is made of how many samples might be required to achieve the desired level of uncertainty in the model. This guess is based on previously generated surrogate models with similar design spaces and is used to set the initial expected sample budget as well as the maximum sample budget. The adaptive sampling continues as normal, but would break if one of the following conditions was met:

1. The RMSE drops below the desired uncertainty.
2. The predicted point at which the uncertainty is reached is greater than twice the maximum value.
3. The maximum number of samples is reached.

This process requires a means of predicting the number of samples that the generator might need based on the developing RMSE history. It was found that a power law + constant \( (ax^b + c) \) consistently performed the best across all four test cases at matching the RMSE history as well as predicting the ultimate sample budget. Table 3.5 shows the \( R^2 \) goodness-of-fit of each of the curves tested to the RMSE history for the four test cases in Figure 3.15.

![Figure 3.15: Comparison of the RMSE history for the four test cases up to a sample budget of 500 and assuming an initial 'guess' of 100](image)
This procedure was tested on all four test cases with the desired $\text{RMSE}$ value of less than $1 \times 10^{-3}$ and using an initial guess of 100 samples (this implied an initial sample of 33 and an absolute maximum of 1000). Table 3.6 shows the number of samples at the point the generation process was terminated, the uncertainty values for the initial 100 sample guess and the uncertainty values at the point the process stopped as well as the reason for stopping.

As can be seen from the table, by easing the restrictions on the budget, all but one case was able to achieve the desired uncertainty in the model. Aero-1 was the only test case not to achieve the required level of uncertainty. It was observed that the model had begun to converge on a value above 1500. The issue is whether the required uncertainty is worth the extensive simulation time. The testing performed here used panel methods to generate the data which runs incredibly quickly. Using DSMC simulations to gather the data would take significantly longer (estimated at least 7 days, assuming 1 hour simulation time and 8 threads). Breaking the simulation at the point it did is therefore justifiable, as it has pushed the uncertainty down, but recognized that the goal was not achievable given the constraints.

As can be seen from the table, in all cases, the initial guess of 100 samples was insufficient to obtain the desired uncertainty of less than $1 \times 10^{-3}$. By easing the restrictions on the budget, all but one case was able to achieve the desired uncertainty in the model (Aero-1 being the only test case not to achieve the required level). A key issue here is that the more the guess is underestimated, the more rounds of adaptive sampling are needed to reach the desired level of uncertainty. The Droplet function, for instance, required 688% more samples to achieve the required uncertainty. All of these samples were added using adaptive sampling, which is a long way from the recommended two-thirds of the total sampling budget. Better predictions based on previous model generations can reduce the dependence on the adaptive sampling acting merely to fill the design space rather than optimize the sample selection. However, as can be seen from Table 3.6, even functions from the same family (such as Aero-1 and Aero-2) required a significantly different number of samples in the final budgets and stopped generating for different reasons. This means that some form of budget easing is always going to be necessary.

<table>
<thead>
<tr>
<th>Test Case</th>
<th>No. Samples</th>
<th>Uncertainty after 100 Samples</th>
<th>Uncertainty at breakpoint</th>
<th>Reason for termination$^1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Droplet</td>
<td>788</td>
<td>$8.89 \times 10^{-3}$</td>
<td>$4.68 \times 10^{-1}$</td>
<td>1</td>
</tr>
<tr>
<td>Franke</td>
<td>260</td>
<td>$5.01 \times 10^{-3}$</td>
<td>$2.20 \times 10^{-2}$</td>
<td>1</td>
</tr>
<tr>
<td>Aero-1</td>
<td>804</td>
<td>$1.03 \times 10^{-1}$</td>
<td>$7.59 \times 10^{-1}$</td>
<td>2</td>
</tr>
<tr>
<td>Aero-2</td>
<td>124</td>
<td>$1.52 \times 10^{-3}$</td>
<td>$7.27 \times 10^{-3}$</td>
<td>1</td>
</tr>
</tbody>
</table>

$^1$ See list of break conditions
3.5 DSMC and Final Evaluation

3.5.1 Introduction

The final stage is to verify that similar results can be achieved using DSMC simulations for the Aero-1 and Aero-2 test cases, as achieved using the panel methods. DSMC is a stochastic method, thus individual sample points will have an additional random error associated with them by default. This error is usually minimized by time-averaging the result of the simulations, thus the longer the simulation runs the more the surface forces converge to a result. However, as was described in Section 3.4.2 due to the poor conditioning of the system matrix $R$, even relatively small errors in the data can have a large impact on the quality of the surrogate model. It is therefore essential to generate a surrogate model using a DSMC source to verify the process still works.

3.5.2 Method

The evaluation was performed by generating a surrogate model for both the Aero-1 and Aero-2 test cases using both the analytical and DSMC simulations to calculate the coefficients of drag of each test case. As previously discussed, the analytical simulations used the linearized equations for rarefied gases developed by Sentman while the DSMC method uses Sandia National Laboratories Rarefied gas code SPARTA.

Both test cases were initiated with an initial sample budget of 100 sample points. This gave an initial sample size of 33, generated using Optimized Latin Hypercube Sampling. Further samples were added in batches (equal in size to the number of available cores on the system) and selected by using the adaptive criteria outlined in Section 3.3.2. The DSMC simulations for the new samples in each batch were then run concurrently on individual cores before regenerating the surrogate model.

As described in Section 3.4.5 the sample budget was relaxed to achieve an uncertainty of $1 \times 10^{-3}$. During the generation process, the LOOCPV uncertainty measure was used to generate the error history for the termination condition. This was also used for the final uncertainty measure as compared to the evaluation grid used in previous sections.

3.5.3 Results & Discussion

Table 3.7 shows the final RMSE for the generated models as well as the total number of samples used and the total run time. Figures 3.16 & 3.17 show the resulting coefficient of drag surfaces for each model using an analytical data source (Figures 3.16a & 3.17a) and a DSMC data source (Figures 3.16b & 3.17b). As can be seen from the graphs in both Figures 3.16 & 3.17 the surface generated by the surrogate model smoothly links all of the samples.

As in previous sections, for both data sources, the Aero-1 test case does not achieve the desired uncertainty of less than $1 \times 10^{-3}$ with the generation process terminating early as a result of the extrapolation criteria (see Section 3.4.5). As discussed previously, this may be as a result of the high contrast between the severe slope at one of the boundaries and the relatively flat nature of the rest of the domain.
CHAPTER 3. GENERATION AND VALIDATION OF A SURROGATE MODEL FOR RAREFIED GAS MODELLING

Table 3.7: Comparison of the RMSE, sample budget and run times for the Analytical and DSMC data sources for the Aero-1 and Aero-2 test cases

<table>
<thead>
<tr>
<th>Test Case</th>
<th>Initial Budget</th>
<th>Analytical</th>
<th></th>
<th>DSMC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>Samples</td>
<td>Run Time</td>
<td>RMSE</td>
</tr>
<tr>
<td>Aero-1</td>
<td>100</td>
<td>$1.24 \times 10^{-2}$</td>
<td>377</td>
<td>00:08:57</td>
</tr>
<tr>
<td>Aero-2</td>
<td>100</td>
<td>$9.76 \times 10^{-4}$</td>
<td>161</td>
<td>00:02:14</td>
</tr>
</tbody>
</table>

The Aero-1 test case was included in the set of test cases as an example of an awkward test case that might be encountered. It is more likely that most of the simulations will be similar in form to the Aero-2 test case, which as can be seen from Table 3.7 did achieve the desired error for both analytical and DSMC simulations. As can be seen from Figures 3.17a & 3.17b the surrogate model is reliably able to reproduce the surface of the design space for both data sources. This shows that the process developed in this chapter is robust enough to handle both the predictable analytical method as well as the stochastic DSMC method.

A considerable concern while generating the surrogate model for the DSMC simulations was the potential for noise in the data to affect the accuracy of the model due to the poor conditioning of the system matrix. As discussed in Section 2.5.3, the DSMC simulations set up for this process (using SPARTA) have an uncertainty in their $C_D$ values of around $1 \times 10^{-3}$. While lower than the target for the surrogate model, the noise this uncertainty represents could still cause difficulty, especially if two sample points happen to be close together. However, as can be seen from Table 3.7, the Aero-2 test case did indeed achieve the desired error threshold when using DSMC to calculate the drag and with a similar number of samples to the analytical case. In general, the DSMC method worked broadly as well as the panel methods as a data source for this surrogate model generation process.

![Figure 3.16: Surrogate Models for test case Aero-1 using the analytical rarefied gas equations developed by Sentman and the DSMC code SPARTA](a) Analytical simulation (b) DSMC simulation}
3.6 Conclusion

(a) Analytical simulation
(b) DSMC simulation

Figure 3.17: Surrogate Models for test case Aero-2 using the analytical rarefied gas equations developed by Sentman and the DSMC code SPARTA

3.5.4 Summary

In this section, the surrogate model was applied to a DSMC data source for the Aero-1 and Aero-2 test cases and compared to an equivalent panel method data source. It was demonstrated that the surrogate models generated using a DSMC data source performed as well as the surrogate model generation that used a panel method as a data source. The surrogate model generation process can therefore be used to generate models for DSMC simulations.

3.6 Conclusion

This chapter has presented the methodology and framework for generating a surrogate model as it was used in this research. It outlined the three primary components of the surrogate model, that being the sampling strategy, the interpolation method and the ‘black box’ function to be approximated as well as a few supporting components. It was shown how an adaptive strategy based on the work of Mackman et al. could provide a better uncertainty and sample budget than simple single-stage designs. It was also shown that for the model used here, a global support radius equal to the length of the longest diagonal in the normalized design hyperspace was sufficient to minimize the uncertainty in the model while ensuring the conditioning of the system matrix was sufficient. It was also seen that a polynomial in this context was not necessary.

To refine the method, a means of relaxing the total sample budget was presented, such that the generation process could break early if the error threshold was met or add additional points if the threshold could be met within a set tolerance of the original Budget. Finally, it was shown how the model generation process worked as well with the DSMC simulations as it did with the panel methods used for the aerodynamic test cases. The method is, therefore, ready to be deployed to fully explore the satellite design space using a DSMC based method.
4.1 Introduction

In the previous chapter, the methodology and framework for generating a surrogate model were defined and verified for use with the aerodynamic simulations using either panel methods or DSMC. In this chapter, the surrogate model will be used to assess the effectiveness of altering the geometry of the satellite to reduce its drag, in an effort to either prolong its operational life or increase its payload capacity.

This chapter will first outline the general methodology used including the aeroshell profiles considered (Section 4.2.3) and how the surrogate model described in Chapter 3 will be implemented (Section 4.2.4). A sensitivity analysis will be performed in Section 4.3 on the elements that describe each of the profiles to determine what effect they have on the drag that they experience in a simple circular orbit. An important driver for this work is to minimize the drag while maximizing the internal volume of the profiles. In order to examine this a series of multi-objective optimizations were performed on the profiles. In Section 4.4 the optimization is performed on a set of profiles with fixed external dimension constraints. This is then extended in Section 4.5 to optimize the aeroshell profiles for a given volume. Work carried out as part of this chapter formed part of an Acta Astronautica paper published in January 2021[7].
4.2 General Methodology

4.2.1 Workflow

The work presented in this chapter focuses on how the geometry of the satellite can be altered to minimize the aerodynamic drag it experiences in VLEO. Reducing the drag would help to save fuel mass or extend the system's operational life, however, since the geometry is being altered this must not limit the internal volume of the satellite. An overview of the workflow for this chapter is provided in Figure 4.1 while Figure 4.2 provides an overview of the functional flow.

As can be seen on the right of Figure 4.1, to explore how the geometry affects the drag it was first necessary to identify what profiles are to be used in the analysis. This required performing a brief review of typical satellite morphologies, including dimensions, of systems that have previously been launched into LEO, thereby informing the choice of dimensions for the design space as is shown in Section 4.2.3.

The first step in the analysis is to identify how the parameters that define each profile affect the atmospheric drag. This will identify whether the feature is beneficial to reducing the drag or merely provides additional volume. To keep the analysis within the bounds of computational complexity, this will be performed for a set of fixed external dimensions, with aspect ratios that could be representative of platforms in the Microsatellite category as outlined in Section 4.2.3.3. The full method and results are shown in Section 4.3.

The next stage of the analysis is to optimize the geometries given a set of scenarios and constraints. Due to the stochastic nature of DSMC simulations, a global optimization strategy was required to effectively interrogate the surrogate models. As will be discussed in Section 4.2.5 Particle Swarm Optimization (PSO) was chosen, as it can traverse large multi-variable problems quickly. Section 4.2.5 will also outline the general constraints placed on the surrogate models while they are being interrogated, as well as the cost functions for the scenarios below.

The first scenario involves performing a multi-objective optimization of the profiles with the aim of minimizing atmospheric drag while maximizing the volume of the satellite. This will be achieved by finding the Pareto-optimum fronts for each of the profiles using the method of linear weightings. To help structure the analysis, the optimization will be performed for a set of fixed external dimensions with different aspect ratios that would be representative of platforms in the Microsatellite category as outlined in Section 4.2.3.3. This optimization scenario, along with the resulting profiles and their drag and internal volumes, is described in Section 4.4.

The second scenario involves obtaining the profiles with the best performance when the desired volume is prescribed. In this instance the optimizer will not be constrained by the five body dimensions. However, to prevent the optimizer from simply seeking the boundary of the data set, the length and width will be constrained in turn. The results of this fixed volume optimization will be presented in Section 4.5.
4.2. GENERAL METHODOLOGY

Figure 4.1: Overview of workflow for Chapter
CHAPTER 4. THE OPTIMIZATION OF SATELLITE'S AERODYNAMICS IN VLEO

Figure 4.2: Overview of functional flow for Chapter 4
4.2.2 Properties of the Fluid

VLEO sits within the lower portion of the Earth’s thermosphere. As discussed in Section 2.3, the air density in this region is higher than in more traditional LEO regions and therefore has the effect of limiting the orbital life of satellites to the order of weeks or days without active drag compensation. There are multiple models available to approximate the fluid properties of this region, namely the density, temperature and chemical composition. For the work performed here the atmospheric model, NRLMSISE-00 was used, following the guideline recommended by the European Space Agency Standard ECSS-E-ST-10-04C[41]. This is an empirical model of the Earth’s atmosphere produced by the US National Research Laboratory and extends from the sea level to the upper limit of the thermosphere at 1000km above local sea level.

The NRLMSISE-00 model has several inputs including time, the location (usually expressed as the altitude, latitude and longitude) as well as recent solar activity and terrestrial magnetosphere strength. The number of inputs reflects the significant variability of the atmosphere and represents a level of complexity that is not needed for this current study. To simplify, a global average of the atmospheric properties were taken at the desired altitude of 200km during high solar activity on the 1st March 2000 as described in Section 2.3.4. This date was chosen to be close to the solar maximum of solar cycle 23, a recent active solar cycle, and thus provided a worst-case scenario for the conditions at the desired altitude. The gas species included in the model were: Oxygen ($O_2$), Nitrogen ($N_2$), Atomic Oxygen ($O_1$), Atomic Nitrogen ($N_1$), Argon ($Ar$), Helium ($He$) and Hydrogen ($H$). A summary of the derived properties used is presented in Table 4.1.

Table 4.1: Properties of the atmosphere at 200km on the 1st March 2000 at 00:00:00 UTC as given by the spherically weighted mean of NRLMSISE-00 model[30] using sampling points at 1° intervals in latitude and longitude

<table>
<thead>
<tr>
<th>Mission Profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude</td>
</tr>
<tr>
<td>Date</td>
</tr>
<tr>
<td>Time</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bulk Fluid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number Density</td>
</tr>
<tr>
<td>Temperature</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atomic Oxygen ($O_1$)</td>
</tr>
<tr>
<td>Nitrogen ($N_2$)</td>
</tr>
<tr>
<td>Oxygen ($O_2$)</td>
</tr>
<tr>
<td>Helium ($He$)</td>
</tr>
<tr>
<td>Atomic Nitrogen ($N_1$)</td>
</tr>
<tr>
<td>Argon ($Ar$)</td>
</tr>
<tr>
<td>Hydrogen ($H$)</td>
</tr>
</tbody>
</table>
4.2.3 Satellite Geometry

4.2.3.1 Introduction

This section will focus on the aeroshell geometry and outline candidate profiles. Drawing on the experience of other flow regimes, there are many examples of aeroshell profiles that could be used here from the simple symmetric cone to the curved Sears-Haack bodies. However, while the Mach number for the flow is high, the $K_n$ is also high and thus the flow can be considered free molecular. This means the flow is also non-viscous and, on the scales being examined, shock waves are not a consideration. For this reason, it is not clear if rounded or blunted bodies would provide any advantage in terms of reducing the drag load on the satellite. However, it would be remiss to dismiss them outright as they may provide a volume advantage over sharper profiles.

Four candidate profiles were identified for examination: a Blunted Wedge (Figure 4.3a, Section 4.2.3.4), an Elliptical Profile (Figure 4.3b, Section 4.2.3.5), a Double-Conic Nose (Figure 4.3c, Section 4.2.3.6) and a Rounded-Conic Nose (Figure 4.3d, Section 4.2.3.7). The first two will explore the effect of flat and rounded nose and tail profiles have on the atmospheric drag while the latter two are variants on the flat plate nose profile. These 4 profiles along with the general assumptions (Section 4.2.3.2), surface properties (Section 4.2.3.8) and satellite dimensions (Section 4.2.3.3) are described below.

![Figure 4.3: Example candidate profiles to be examined in this section](image-url)
4.2.3.2 General Assumptions

The analysis will focus on the central body of the satellite, ignoring the solar panels or any instruments that are externally mounted. This is because the size and impact the solar panels have on the drag are highly dependent on the overall satellite configuration and the propulsion system that is implemented and therefore difficult to generalize. The initial central body will be assumed to be a cuboid with smooth surfaces and thus lacking any extrusion, recesses or holes. The surface properties used in the simulations will be discussed more fully in Section 4.2.3.8. Given these assumptions, the analysis was reduced to a two-dimensional axisymmetric body which improves the computation time per sample. It has been shown in previous work that the smaller the frontal area, the lower the atmospheric drag [27], so the bodies were orientated such that their long axis was parallel to the velocity vector.

4.2.3.3 Dimensions

A targeted study was performed to determine typical dimensions for the central body of small spacecraft. Data from 179 satellites was compiled using sources from the original manufacturer and supplemented with data from Encyclopaedia Astronautica[89] and the Union of Concerned Scientists (UCS)[90]. The results of this study are summarized in Table 4.2 for nanosatellites, microsatellites and minisatellites. Previous work [12, 13, 2] has shown that satellites in the Microsatellite category (10-100kg) are commercially feasible in VLEO, so will be the focus of this line of work. This is significantly smaller than either GOCE or SLATS, which were small and minisatellites respectively, but presents a good starting point for the analysis. To fully explore the possible body geometries for microsatellites, each body profile will be generated with a width \((w_b)\) range of 0.2 - 0.75 m and a body length \((l_b)\) range 0.2 - 1.5 m.

Table 4.2: Typical dimensions for each satellite category used in the simulations

<table>
<thead>
<tr>
<th>Category</th>
<th>Mass Range [kg]</th>
<th>Body Length Range* [m]</th>
<th>Body Width Range* [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nanosatellite</td>
<td>1 - 10</td>
<td>0.1 - 0.5</td>
<td>0.1 - 0.3</td>
</tr>
<tr>
<td>Microsatellite</td>
<td>10 - 100</td>
<td>0.2 - 1.5</td>
<td>0.2 - 0.75</td>
</tr>
<tr>
<td>Minisatellite</td>
<td>100 - 500</td>
<td>0.5 - 5</td>
<td>0.5 - 2.5</td>
</tr>
</tbody>
</table>

* Where possible satellite dimensions were sourced from the original manufacturer and supplemented with data from Encyclopaedia Astronautica[89] and the UCS[90].

4.2.3.4 Blunted Wedge

The most basic form of nose cone geometry is that of a symmetric cone. The flat plates make this kind of profile easier to manufacture, though the diagonal nature can impact the available usable internal space. A simple variation on this is to flatten the nose resulting in a blunted cone. It is likely that blunting the cone may not help improve the drag but this research aims to find profiles that are practical and blunting may help to improve the usable volume to drag ratio.
Another factor that will be considered with this profile is the effect that the rear of the satellite has on the drag. In this instance, this will also be a blunted symmetric cone which, along with the nose geometry, describes a blunted diamond-wedge body as shown in Figure 4.4. This is similar to the symmetric candidates tested in Chapter 2.

It is not clear what the impact of varying the rear geometry of the satellite will be since objects in a rarefied gas will normally just leave a void that slowly refills behind them. However, it was highlighted in Section 2.5.2 that the surfaces along the side of the satellite do have a reasonable contribution to the overall drag. It was therefore of interest to see whether altering these surfaces would have an impact on the drag and help reduce it.

The Blunted Wedge profile is defined by 4 parameters: the nose length \( l_n \), nose radius \( r_n \), tail length \( l_t \) and tail radius \( r_t \). The ranges for these parameters as used in this chapter are given in Table 4.3. Given the definition of the parameters, it would be possible for the nose and tail profiles to overlap, so to prevent unrealistic geometries, \( l_n \) will take precedence. This means that if \( l_n \) plus \( l_t \) would be larger than the body of the satellite, the tail is truncated to fit the body.

Figure 4.4: Diagram showing the general form of the Blunted Wedge with key dimensions and direction of the velocity vector

Table 4.3: Parameters of the Blunted Wedge used in this chapter

<table>
<thead>
<tr>
<th>parameter</th>
<th>min</th>
<th>max</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>nose radius</td>
<td>0.0</td>
<td>0.5</td>
<td>proportion of the body width</td>
</tr>
<tr>
<td>nose length</td>
<td>0.0</td>
<td>1.0</td>
<td>proportion of the body length</td>
</tr>
<tr>
<td>tail radius</td>
<td>0.0</td>
<td>0.5</td>
<td>proportion of the body width</td>
</tr>
<tr>
<td>tail length</td>
<td>0.0</td>
<td>1.0</td>
<td>proportion of the body length*</td>
</tr>
</tbody>
</table>

*As the tail and nose will overlap, the nose profile takes precedence
4.2.3.5 Elliptical Profile

An alternative to using the flat plated symmetric cone is to employ a form of a curved profile. Given the rarefied, non-viscous nature of the flow at the altitudes of interest, it is not certain that a rounded profile such as this would provide any significant advantage with regard to drag when compared to a Blunted Wedge. Where this profile may have the advantage is in achieving a better drag to volume ratio, but this will need to be explored.

In this instance, a simple Elliptical profile was chosen for the nose and tail, which both smoothly meet the non-curved parallel walls of the central section, as shown in Figure 4.5. Given that it is uncertain whether rounded surfaces would provide any benefit in this flow regime, more complex curved profiles such as parabolic or ogive have not been considered here but could be the basis of future work.

The Elliptical profile is defined by 2 parameters: the nose length \( l_n \) and the tail length \( l_t \). The range for these parameters is given in Table 4.3. In a similar manner to the Blunted Wedge, it would be possible for the nose and tail profiles to overlap, so to prevent unrealistic geometries \( l_n \) will take precedence. This means that if \( l_n \) plus \( l_t \) is larger than the body of the satellite, the tail is truncated to fit the body. Additionally, most implementations of the DSMC method are not able to accept curved surfaces as definitions for the geometry so the rounded surface will be approximated as a series of short faces. It was assumed that the rounded segments were approximate by 20 short flat faces as this provided a good compromise between speed and accuracy.

![Figure 4.5: Diagram showing the general form of the Elliptical profile with key dimensions and direction of the velocity vector](image)

Table 4.4: Parameters of the Elliptical profile used in this chapter

<table>
<thead>
<tr>
<th>parameter</th>
<th>min</th>
<th>max</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>nose length ( l_n )</td>
<td>0.0</td>
<td>1.0</td>
<td>proportion of the body length</td>
</tr>
<tr>
<td>tail length ( l_t )</td>
<td>0.0</td>
<td>1.0</td>
<td>proportion of the body length*</td>
</tr>
</tbody>
</table>

*As the tail and nose will overlap, the nose profile takes precedence
4.2.3.6 Double-Conic Nose

In the rarefied atmosphere of VLEO, it is expected that the blunted nose geometry will likely result in a higher level of drag on the body. In an attempt to mitigate this, one solution is to angle the blunted nose away from the flow, resulting in the Double-Conic Nose, as shown in Figure 4.6. Unlike the previous two profiles, this example will simply look at the nose profile, as the aim is to see if this provides an advantage over the Blunted Wedge, particularly when the volume of the object is fixed.

The profile is defined by the total nose length \( l_n \) as well as the front-nose length \( l_f \) and front-nose radius \( r_n \) for the additional cone. The ranges for these parameters are given in Table 4.5. Given the definition of the parameters, it would be possible for the nose to become concave, so two constraints are included. Firstly, \( l_f \) is always shorter than \( l_n \). Secondly, the half cone angle of the front nose must be greater than or equal to the half-cone angle of the second cone.

![Figure 4.6: Diagram showing the general form of the Double-Conic nose with key dimensions and direction of the velocity vector](image)

<table>
<thead>
<tr>
<th>Table 4.5: Parameters of the Double-Conic nose used in this chapter</th>
</tr>
</thead>
<tbody>
<tr>
<td>parameter</td>
</tr>
<tr>
<td>total nose length</td>
</tr>
<tr>
<td>front cone length</td>
</tr>
<tr>
<td>front cone radius</td>
</tr>
</tbody>
</table>

\( *l_f \) must always be shorter than \( l_n \).
4.2. GENERAL METHODOLOGY

4.2.3.7 Rounded-Conic Nose

An alternative method for potentially reducing the drag on the blunted nose of the wedge is to round out the nose as shown in Figure 4.7. Given the rarefied, non-viscous nature of the flow, it is not expected that the rounding of the nose will help to guide the flow around the profile. However, it can be expected that the pressure on the rounded surfaces will gradually reduce till the tapered section is reached, which should allow the profile to achieve a similar drag value as the Blunted Wedge, but with greater internal volume, in much the same way as the Elliptical profile. As with the Double-Conic profile, this body will simply look at the nose profile, as the aim is to see if this provides an advantage over the Blunted Wedge. The Rounded-Conic nose profile is defined by 2 parameters: the nose length ($l_n$) and the nose radius ($r_n$). The ranges for these parameters are given in Table 4.6. The only restriction on the parameters, in this case, is that $r_n$ must be smaller than $l_n$.

![Diagram showing the general form of the Rounded-Conic nose with key dimensions and direction of the velocity vector](image)

Figure 4.7: Diagram showing the general form of the Rounded-Conic nose with key dimensions and direction of the velocity vector

<table>
<thead>
<tr>
<th>parameter</th>
<th>min</th>
<th>max</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>nose radius</td>
<td>0.0</td>
<td>0.5</td>
<td>proportion of the body width</td>
</tr>
<tr>
<td>nose length</td>
<td>0.0</td>
<td>1.0</td>
<td>proportion of the body length</td>
</tr>
</tbody>
</table>
4.2.3.8 Surface Properties

An important aspect of modelling rarefied gases is the interaction between the particle and the surfaces of the test body. The DSMC tool SPARTA uses a diffuse model to replicate the particle-surface interactions within a rarefied gas. In this diffuse model, these interactions are governed by the surface temperature and the energy accommodation coefficient ($\alpha_T$) \[48\]. The $\alpha_T$ describes how much energy the incoming particle loses to the surface and is defined as

$$\alpha_T = \frac{E_i - E_r}{E_i - E_w}$$

where $E_i$ is the kinetic energy of the incident molecule, $E_r$ is the kinetic energy of the re-emitted molecule, and $E_w$ is the kinetic energy the re-emitted molecule would have if it were re-emitted from the surface at the surface’s temperature. When $\alpha_T$ equals 1, complete thermal accommodation has occurred, i.e. the particle has been re-emitted at the same temperature as the impacted surface and therefore has lost some energy to the surface.

At present, there is some uncertainty around the value of $\alpha_T$ for materials in VLEO. Empirical findings from satellites in orbit have shown that $\alpha_T$ ranges in value from 0.86 to 1, with limited dependences on the underlying material\[27, 32\]. This contrasts, however, with early ground testing, which has shown some dependence on the material choice, though often these are not comparable conditions to those seen in orbit. While it is likely that many materials will still exhibit mostly diffuse or quasi specular behaviour, this is not known with any great certainty. Recent work by Roberts et al. \[22, 33\] seeks to improve this understanding by examining the gas-surface interactions between the material and the atmosphere under the conditions observed in orbit. At the time of writing, the Rarefied Orbital Aerodynamics Research facility (ROAR) is still being commissioned, but once operational will be capable of simulating the free molecular flow and levels of atomic oxygen flux observed in VLEO.

For the work presented here an energy accommodation coefficient $\alpha_T = 0.95$ was selected, representing a mostly diffuse environment with some specular reflections occurring. Future work could explore if there is any dependence between the selected profile and $\alpha_T$, however, it is known that as surfaces degrade in the atomic oxygen-rich environment of VLEO, the surface interactions tend to be more diffuse as the satellite ages. The surface temperature of the satellites is taken to be $300 \text{K}$, which is broadly within the operational range of most terrestrial satellites. These values were used to define the properties of the surface for both the Panel Methods and the DSMC simulations.

4.2.4 Surrogate Model

4.2.4.1 The Model

A surrogate model was generated for the design space outlined in Section 4.2.3 using the method outlined in Chapter 3 for each profile. The design space for each profile consisted of the length and width of the body ($l_b$ & $w_b$ respectively) plus the parameters that define each profile. The outputs of the surrogate model are the drag force, the coefficient of drag ($C_D$) and the internal volume of the geometry represented by the evaluation point. The $C_D$ for each sample point was calculated using the DSMC software SPARTA as described in Section 2.4.3.5.
The primary reason for adopting a surrogate-based approach is to allow the use of DSMC methods to help explore each aeroshell profile. As seen in Section 2.5.3, DSMC simulations are stochastic methods that often require a significant computational resource to converge to a result. This makes it difficult to explore a design domain using methods that rely solely on directly simulating each evaluation point. This limits the number of optimization scenarios that can be examined. A surrogate model (as described in Chapter 3) can approximate the ‘black box’ function of the DSMC results for a given profile using an optimized set of samples. The optimizer can then interrogate the surrogate model which is trivial when compared to running a full DSMC simulation for each evaluation. This means that several different optimization scenarios and constraints can be tested on the same surrogate data set, thus reducing the number of DSMC simulations required to explore the design space.

For the training data sets used in this work, it was estimated that around 2000 samples would be required. This was based on observations from previously generated surrogate models while validating the process. As outlined in Section 3.3.2, this gave an initial sample size of 667 generated using OLHS. Further samples were added in batches (equal in size to the number of available cores on the system) and selected using the adaptive criteria outlined in Section 3.3.2. The DSMC simulations for the new samples in each batch were then run concurrently on individual cores, before regenerating the Surrogate Model. While the maximum number of samples was set as 2000, this condition was then relaxed to achieve a lower RMSE and data collection was terminated when the system determined the RMSE could not be improved upon, as described in Section 3.4.5.

The range of the $C_D$ is expected to be between 2 and 3 with much of the detail in the 1st and 2nd decimal place. For this reason, an RMSE of less than $1 \times 10^{-3}$ is required to confidently capture the detail of the sample space. For these data sets the RMSE was calculated using the LOOCPV method as described in Section 3.3.4.2. The RMSE history of the $C_D$ as more samples are added is shown in Figure 4.8.

As seen in both Figure 4.8 and Table 4.7 for the Blunted Wedge profile the final RMSE is $9.54 \times 10^{-4}$, with a 95% confidence bracket of $2.86 \times 10^{-3}$. Thus, for a 95% confidence interval, this implies the model $C_D$ is accurate to within $\pm 0.00286$ of the DSMC simulations. In some circumstances, greater accuracy can be achieved by increasing the number of sampling points. However, this presents the risk of over-fitting the problem, which may ultimately increase the error margin. In the case of this data set, the generator terminated sampling, as it determined no further improvement in the accuracy was possible.

<table>
<thead>
<tr>
<th>Initial number of Sample Points</th>
<th>667</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Sample Points</td>
<td>6827</td>
</tr>
<tr>
<td>Number of Evaluation Points</td>
<td>500</td>
</tr>
<tr>
<td>Final RMSE</td>
<td>$9.54 \times 10^{-4}$</td>
</tr>
<tr>
<td>95% Confidence Bracket</td>
<td>$2.86 \times 10^{-3}$</td>
</tr>
</tbody>
</table>
Figure 4.8: Error history of the surrogate model data set used in this work for the Blunted Wedge profile

Slices through the Blunted Wedge design space are presented in Figures 4.9 & 4.10. Figure 4.9 shows how the $C_D$ and internal volume varies for $l_n$ and $r_n$ while the tail is held in a blunt configuration ($l_t$ and $r_t$ both equal to 0). Similarly, Figure 4.10 shows these variations for $l_t$ and $r_t$ while the nose is held in a blunt configuration ($l_n$ and $r_n$ both equal to 0).

4.2.4.2 Speed of Generation

One of the reasons for using a smarter sampling technique like adaptive sampling is that, for a given uncertainty in the model, it should require fewer samples than simpler FF and OLHS methods. This means it should generate the surrogate model much quicker, however, being a smarter method also comes with additional computational overheads.

New samples were normally added in batches equal to the number of computer cores being used at the time (in this case 16). This meant that the coefficients of the surrogate model, required to calculate the Laplacian for the adaptive criteria in Equation 3.19 only had to be calculated once per batch. However, the separation function, which itself was a surrogate model (see Section 3.3.2.4) had to be recalculated with each new sample point. This required the inversion of a system matrix which for $N$ samples has a complexity of $O(N^3)$.

For the Blunted Wedge profile with 6 parameters the maximum time spent evaluating the adaptive criteria was 1 minute and 5 seconds per sample (17 minutes and 24 seconds for a batch of 16). While this is long, it is still significantly shorter than the 1 hour 42 minutes required to simulate the body. Additionally, this occurred during the last adaptive evaluation which had the largest matrix to invert (6827 samples). Therefore, the generation process remained dominated by the samples themselves and not the process to find them.
4.2. GENERAL METHODOLOGY

Figure 4.9: A slice through the surrogate model hyperspace, showing the variation of the $C_D$ with nose geometry with fixed tail geometry for the Blunted Wedge profile

Figure 4.10: A slice through the surrogate model hyperspace, showing the variation of the $C_D$ with tail geometry with fixed nose geometry for the Blunted Wedge profile
Table 4.8 provides a comparison of the performance of the generator using adaptive sampling against a simpler sample scheme such as FF for the Blunted Wedge profile with 2, 4, and 6 active design parameters under similar atmospheric conditions (Table 4.1). For the adaptive sampling, the table shows the number of samples (N) needed by the surrogate model generator, as well as the total run time, to achieve the desired uncertainty of $1 \times 10^{-3}$ in the $C_D$ (calculated using LOOCPV). Also shown in Table 4.8 is the sample density along each parameter. This is simply taken to be the $\sqrt{N}$ where $n$ is the number of parameters and $N$ is the total number of samples.

As observed in Section 3.4.4, a gridded system requires significantly more samples to achieve the same level of uncertainty in the surrogate model, but the process to determine the sample sites was much quicker. To provide a comparison, an approximate lower bound on the number of samples required in a gridded scheme is shown in Table 4.8 and was calculated assuming the same sample density per parameter as the adaptive sampling model with 2 parameters.

As can be seen from Table 4.8, a gridded approach to the Blunted Wedge with 6 parameters would have required at least 4,826,809 samples to achieve the same level of uncertainty in the model as the adaptive approach. This is $707 \times$ more samples than were needed under the adaptive sampling used here. If this were the case, the DSMC simulations alone would take about 44 years using 16 cores. While this run time could be improved by running the DSMC simulations over more cores, approximately 10725 cores would be needed to complete the simulations in the same time frame as the adaptive sampling method.

Table 4.8: Approximate number of samples needed by the adaptive sampling process used for the Blunted Wedge under similar atmospheric conditions but with varying number of parameters Note: assumes 16 cores

<table>
<thead>
<tr>
<th>$n$</th>
<th>Adaptive</th>
<th></th>
<th>Gridded</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
<td>$\sqrt{N}$</td>
<td>Run Time</td>
<td>$N$</td>
</tr>
<tr>
<td>2</td>
<td>177</td>
<td>13</td>
<td>17:07:44</td>
<td>169</td>
</tr>
<tr>
<td>4</td>
<td>2100</td>
<td>6.76</td>
<td>9 days</td>
<td>28,561</td>
</tr>
<tr>
<td>6</td>
<td>6827</td>
<td>4.36</td>
<td>24 days</td>
<td>4,826,809</td>
</tr>
</tbody>
</table>

4.2.5 Optimization Strategy

4.2.5.1 Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) is a metaheuristic strategy for global optimizations inspired by the collective social behaviour of animals such as swarming insects, flocking birds or schools of fish[91]. PSO consists of a population of particles that traverse the design space, sampling candidate solutions as they move with each iteration. A particle’s movement through the design space is then updated based on the location of its best-known candidate solution, as well as the
best-known candidate of the population as a whole. Eventually, the particles in the swarm converge on the global optimum for that design space, though this is not always guaranteed.

PSO was chosen for this work because it can traverse large multi-variable problems quickly compared to other global optimizers. This is because PSO relies solely on the observed candidate solutions to find the optimal solution and does not attempt to calculate the gradients at each iteration. However, each particle will still need to interrogate the surrogate model, so a balance must still be struck between speed and accuracy.

The work performed here used the MATLAB R2020a implementation of PSO.

4.2.5.2 Parameters, Bounds and General Constraints

The aeroshell profiles that are to be examined as part of this work and the parameters that define them were outlined in Section 4.2.3. These can be summarised into 7 parameters: body length ($l_b$), body width ($w_b$), nose radius ($r_n$), nose length ($l_n$), fore nose length ($l_f$), tail length ($l_t$) and tail radius ($r_t$). These parameters, along with their design boundaries and the profiles they apply to, are outlined in Table 4.9. It should be noted that while the parameters are similar, they may have slightly different effects on the geometry, depending on the profile. For instance, for the Blunted Wedge and Double-Conic nose, $r_n$ defines the width of the blunted section of the nose. In the case of the Rounded-Conic nose, it represents the radius of the rounded tip.

In addition to the bounds described in Table 4.9, there are a few constraints placed on the parameters. for instance, given the definition of parameters $l_n$ and $l_f$, it would be possible for the nose and tail profiles to overlap, so to prevent unrealistic geometries, $l_n$ will take precedence. This means that if $l_n$ plus $l_f$ is larger than the body of the satellite, the tail is then truncated to fit the body. Similarly, given the definition of parameters $r_n$, $l_n$ and $l_f$, it might be possible for the nose of the Double-Conic profile to become concave, so two constraints are included. The first constraint is that $l_f$ is always shorter than $l_n$. The second constraint is that the half cone angle of the front nose must be greater than or equal to the half-cone angle of the second cone.

Table 4.9: Geometric bounds on the design spaces and the profiles they apply to

<table>
<thead>
<tr>
<th></th>
<th>$l_b$</th>
<th>$w_b$</th>
<th>$r_n$</th>
<th>$l_n$</th>
<th>$l_f$</th>
<th>$l_t$</th>
<th>$r_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>min</td>
<td>0.2</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>max</td>
<td>1.5</td>
<td>0.75</td>
<td>0.5</td>
<td>1.0</td>
<td>0.5</td>
<td>1.0</td>
<td>0.5</td>
</tr>
<tr>
<td>Blunted Wedge</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Elliptical</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Double-Conic nose</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Rounded-Conic nose</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
4.2.5.3 Parameter Sensitivity

Before optimizing the whole geometry, it is useful to assess how individual elements of the geometry affect the atmospheric drag ($C_D$) that the profile experiences. It is therefore useful to perform a sensitivity analysis on each of the profiles described in Section 4.2.3. This will be achieved by fixing one of the characteristic parameters of the profile in turn and optimizing the other parameters using the simple cost function

$$ f_c = f_{C_D}(G) \tag{4.2} $$

It is also useful to constrain the design space while analysing the features of each profile. In this case, the bodies have been constrained to discrete aspect ratios with exterior dimensions as follows: $0.25m \times 0.5m$, $0.25m \times 1m$, $0.25m \times 1.5m$, $0.5m \times 1m$, and $0.5m \times 1.5m$. This provided a range of aspect ratios at both ends of the microsatellite scale. This has meant the capabilities of these aerodynamic features can be examined for microsatellites. From a design perspective, these constraints could be considered as a result of internal and external geometric constraints placed on the design: internally this could stem from fixed system dimensions such as a minimum diameter for the payload, while externally constraints may be set by the packing requirements within the fairing. It should be noted that $0.25m \times 0.5m$ and $0.5m \times 1m$ have the same aspect ratio, which is intended as a means of verification. It would be expected that these two bodies would have similar $C_D$ and thus this acts as a check for the data sets.

4.2.5.4 Fixed Body Optimization

As outlined at the beginning of the chapter there are two key objectives: to minimize the drag on the satellite body and to maximize the internal volume of the satellite. To find the Pareto-optimal front for this multi-objective problem, the method of linear weighting was used. This is the simplest method of obtaining the Pareto-optimal fronts for the multi-objective problem and is guaranteed to find solutions on the entire Pareto-optimal set, so long as the front remains convex. Due to its simplicity and ease of implementation, this method was initially used during the development of the surrogate model in Chapter 3 to help analyse the data being generated. As will be seen from the results presented in this Section 4.4, it was found that the Pareto-optimal fronts formed by the data sets are all convex and thus a more complex approach, such as a multi-objective genetic algorithm, would not be necessary for this work.

The cost function ($f_c$) for the method of linear weighting is given by:

$$ f_c = \alpha f_{C_D}(G) - (1 - \alpha) f_V(G) \tag{4.3} $$

The objective function for the atmospheric drag $f_{C_D}(G)$ will take the satellite geometry and return the $C_D$ for the orbital environment by interrogating the surrogate model. The objective function for the internal volume $f_V(G)$ takes the satellite geometry and returns the volume. As the simulations were performed in two dimensions, the area was used to represent the volume. As the outputs of these functions are of a similar order of magnitude, it was not deemed necessary to scale them in this case. The parameter $\alpha$, was a relative weighting of the two objectives, so when $\alpha = 0$, the optimizer maximized the internal volume, while if $\alpha = 1$, the optimizer minimized the drag.
As was the case with the parameter sensitivity, it is also useful to constrain the design space while analysing the features of each profile. In this case, the bodies have been constrained to discrete aspect ratios with exterior dimensions as follows: 0.25m × 0.5m, 0.25m × 1m, 0.25m × 1.5m, 0.5m × 1m, and 0.5m × 1.5m.

4.2.5.5 Fixed Volume Optimization

While the external dimensions can often be a constraining factor on the geometry of a satellite, it is more common to prescribe a desired internal volume for the satellite body. Therefore, in this scenario, the profiles will be optimized given the desired volume.

Unlike the previous scenario, this optimization will aim to minimize the drag force \( f_D(x) \). In the previous scenario the bodies of the satellite were fixed, so it was possible to optimize based on the \( C_D \). This was convenient as the \( C_D \) provides an easy comparison between the various aspect ratios. However, without fixed external dimensions, the reference area can vary with the volume. It is, therefore, possible for two bodies to have a similar \( C_D \), but different reference areas (and thus different drag forces). Since the ultimate objective is to reduce the fuel consumption by reducing the drag force \( F_D(x) \), it is the drag force that needs to be minimized in this instance.

A second consideration is that the PSO, as coded in MATLAB, does not support non-linear constraints and thus any non-linear constraint must be included in the cost function itself. In this instance it is the volume that is to be fixed, so a constraining factor \( C_V \) can be defined:

\[
C_V(G) = \exp(\epsilon(V_{\text{required}} - f_V(G))^2)
\]

(4.4)

\( C_V \) is equal to 1 in regions of the design space that have the desired volume, but is a large positive number everywhere else. The scaler \( \epsilon \) helps to tune the constraint to the original cost function. The cost function for this section therefore becomes:

\[
f_c(G) = f_D(G)C_V(G)
\]

(4.5)

As seen in Section 2.5.2, since the majority of the drag force comes from the forward-facing surfaces, it is likely that the optimizer would opt for the thinnest profile available for the data sets. This is not desirable as such thin profiles might not be of practical use. For this scenario, the profile body’s length and width will be constrained individually between the range of values presented in Table 4.9.
4.3 Drag Sensitivity

4.3.1 Overview

Before optimizing the whole geometry, it is useful to assess how individual elements of the geometry affect the drag that the profile experiences. This section describes a brief sensitivity analysis on each of the profiles described in Section 4.2.3. The simulations that populated the surrogate models used in this section were performed using DSMC and assumed a circular Earth orbit of 200km altitude and an $\alpha$ of 0.95.

Analysis of the Blunted Wedge profile will be divided between the Nose and Tail section (as seen in Sections 4.3.2 & 4.3.3 respectively). Section 4.3.4 will focus on the collective geometry of the Elliptical profile and finally, Section 4.3.5 will look at the ‘Nose-only’ profiles.

4.3.2 Results & Discussion: Blunted Wedge – Nose Profile

This section will focus on the impact that the nose profile of the Blunted Wedge will have on the atmospheric drag on the body ($C_D$). The Blunted Wedge nose profile is described by the nose radius ($r_n$) and the nose length ($l_n$), as described in Section 4.2.3.4. Figures 4.11 & 4.12 show the minimum $C_D$ achievable when $l_n$ and $r_n$ are fixed respectively. In both figures, the other element was allowed to vary to achieve the lowest $C_D$, while the rear of the body was held in a blunt configuration.

The simulations performed for this section were carried out using DSMC and assumed a circular orbit of 200km altitude and $\alpha = 0.95$. Figure 4.13 shows the normalised density field as well as the surface pressure and force vectors around selected configurations of the nose geometry on a 0.5m $\times$ 1.0m body. The density field is normalised to the density of the far-field as given in Table 4.1 (the dotted red line indicates regions of the density field that are equal to the far-field density). The surface pressure and force vectors are to scale but have been magnified by $1 \times 10^4$.

Table 4.10: Unconstrained optimized bodies for minimum $C_D$ for varying nose geometry and fixed (blunt) tail geometry using the Blunted Wedge profile

<table>
<thead>
<tr>
<th>Profile Size (height $\times$ length)</th>
<th>Cuboid $C_D$</th>
<th>Opt. Profile*</th>
<th>Opt. $C_D$</th>
<th>% Red. in Drag</th>
<th>Volume $m^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25m $\times$ 0.5m</td>
<td>2.428</td>
<td></td>
<td>1.932</td>
<td>20.4</td>
<td>0.063</td>
</tr>
<tr>
<td>0.25m $\times$ 1.0m</td>
<td>2.679</td>
<td></td>
<td>1.953</td>
<td>27.1</td>
<td>0.125</td>
</tr>
<tr>
<td>0.25m $\times$ 1.5m</td>
<td>2.930</td>
<td></td>
<td>2.066</td>
<td>29.5</td>
<td>0.188</td>
</tr>
<tr>
<td>0.50m $\times$ 1.0m</td>
<td>2.433</td>
<td></td>
<td>1.933</td>
<td>20.5</td>
<td>0.250</td>
</tr>
<tr>
<td>0.50m $\times$ 1.5m</td>
<td>2.563</td>
<td></td>
<td>1.931</td>
<td>24.6</td>
<td>0.375</td>
</tr>
</tbody>
</table>

* direction of flight right-to-left with particle flow from the left
and $1 \times 10^5$ respectively for clarity. The examples presented in Figure 4.13 are the original cuboid ($r_n = 1, l_n = 0$) (Figure 4.13a), a partially blunted nose profile ($r_n = 0.5, l_n = 1$)(Figure 4.13b) and a fully pointed nose profile ($r_n = 0, l_n = 1$) (Figure 4.13c).

In the case of fixed $r_n$, the lowest $C_D$ was achieved at the maximum $l_n$, in this case, the full length of the body. Similarly, in the case of fixed $l_n$, the minimum $C_D$ was generally achieved when $r_n$ was minimized. The exception to this was when $r_n$ was large or $l_n$ was small respectively, in other words, as the profile approached that of a simple cuboid. This was a limit of the model since, as is seen in Figure 4.9, the lower right boundaries of the graph defined the simple cuboid body, and as the profile approached this boundary there was much less variation in $C_D$ and thus the optimizer was more susceptible to local minima. Critically this was observed while testing using panel methods as well, which implies it is a limitation of the surrogate model with this design space.

The results from Figures 4.11 & 4.12 imply that, in the absence of constraints on the nose geometry, for a blunt tail configuration, the optimum shape should be the sharpest nose profile possible. This is what was observed in Table 4.10 which shows the profile and $C_D$ achieved when optimizing for minimum drag for the different bodies while holding the tail in a blunt configuration. The table also provides the $C_D$ of the unoptimized cuboid, i.e. the untapered body, for comparison, along with the relative reduction in the $C_D$. These results show that a maximum reduction of 20-30% in $C_D$ was achieved when compared to the original cuboid bodies, depending on the aspect ratio.

The results presented so far do call into question the usefulness of including $r_n$ in the profile definition. The radius of the nose acted to blunt the profile, increasing the frontal area, which fully reflected the flow back on itself. However, it also had the effect of decreasing the internal angle of the nose for a given $l_n$. As seen in Figure 4.12, for the lower aspect ratio profiles, there was a modest increase in $C_D$ when the radius of the nose was less than 50% of the width body. For the 1:2 profiles, this increase in $C_D$ is about 7.3% when $r_n = 50\%$ compared to 25.9% when $r_n = 100\%$. There appears to be less of an advantage for bodies with large aspect ratios as the internal angle was already quite small when the nose is as large as it can be and thus has limited variation over the range of nose radii. However, it was observed that if $l_n$ was constrained to increase the internal volume, the $C_D$ curve became similar to that of the lower aspect ratio bodies (though these values were higher due to the increased length of the profile $l_b$). So, while it was not ideal for reducing the drag, it was good for maximizing volume.
Figure 4.11: Optimum $C_D$ for varying nose length ($l_n$) with fixed (blunt) tail geometry using the Blunted Wedge profile

Figure 4.12: Optimum $C_D$ for varying nose radius ($r_n$) with fixed (blunt) tail Geometry using the Blunted Wedge profile
4.3. DRAG SENSITIVITY

(a) Original Cuboid ($r_n = 1, l_n = 0$)

(b) Partially Blunted ($r_n = 0.5, l_n = 1$)

(c) Fully Pointed ($r_n = 0, l_n = 1$)

Figure 4.13: Normalised density field and surface force vectors around the nose geometry of the $0.5m \times 1.0m$ Blunted Wedge profile. The density field is normalised to the free-stream density (red dotted line). Pressure vectors (red) and Force vectors (green) to scale and magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity.
4.3.3 Results & Discussion: Blunted Wedge – Tail Profile

The drag the body experiences is predominantly determined by the body’s frontal projected area. However, due to the random motion of the particles in the gas, the incoming flow is not perfectly parallel, and so there are particle-surface interactions taking place along the side of the body as well. In some cases, this could contribute up to 10% of the atmospheric drag that the body experienced [3] and as can be seen by Table 4.10, the longer the body the higher the $C_D$. While this was not necessarily a large contribution, it was still worth asking whether the rear profile could be altered in any way to reduce its contribution to the atmospheric drag. This section will therefore focus on the impact the tail profile of the Blunted Wedge had on the atmospheric drag the satellite body would experience. In this instance, the tail profile was defined by the tail radius ($r_t$) and the tail length ($l_t$) (as outlined in Section 4.2.3.4).

The simulations performed for this section were carried out using DSMC and they assumed a circular orbit of 200km altitude and an energy accommodation of $\alpha_T = 0.95$. Figure 4.16 shows the normalised density field, as well as the surface pressure and force vectors around selected configurations of the tail geometry on a $0.5m \times 1.0m$ body. The density field is normalised to the density of the far-field as given in Table 4.1 (the dotted red line indicates regions of the density field that are equal to the far-field density). The surface pressure and force vectors are to scale, but have been magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity. The examples presented in Figure 4.16 are the the original cuboid ($r_n = 1, l_n = 0$) (Figure 4.16a), a partially blunted nose profile ($r_n = 0.5, l_n = 1$)(Figure 4.16b) and a fully pointed nose profile ($r_n = 0, l_n = 1$) (Figure 4.16c).

The graphs in Figures 4.14 & 4.15 show the minimum $C_D$ achievable when $r_t$ and $l_t$ are fixed respectively. In both cases, the other element was allowed to vary to achieve the lowest $C_D$, while the front of the body was held in a blunt configuration. In contrast to the front of the satellite and as a direct result of being in a non-molecular flow, the satellite leaves a fluid void directly behind it. This has a much lower density and, as a result, a higher mean free path and $K_n$ than the surrounding fluid. This means the chances of particle-particle interactions taking place is much lower and thus has less of an impact on the predicted $C_D$. This explains why the resulting curves in Figures 4.14 & 4.15 are broadly parallel if slightly offset.

From the results shown in Figures 4.14 & 4.15, it was clear that altering the tail of the satellite would indeed reduce the drag that the body experienced. For instance, when $l_t$ was varied (Figure 4.15), $C_D$ varied proportionally, decreasing as the length increased. On the other hand (Figure 4.14), the $C_D$ appears to have an exponential relationship with $r_t$ when it is fixed. In fact, for profiles with an aspect ratio of 1:2, there is no significant increase in $C_D$ when $r_t \leq 50\%$. Some of the other profiles exhibit similarly plateauing at lower tail radii, though this becomes less pronounced as the aspect ratio increases.

Optimizing the whole tail profile ($r_t$ and $l_t$) for minimum drag while holding the nose in a blunt configuration gave a reduction in $C_D$ of between 10-20%, depending on the aspect ratio. Table 4.11 shows the result of this optimization for the five bodies as well as the profiles. The table also shows the $C_D$ of the unoptimized cuboid, i.e. the untapered body, for comparison along with the reduction in $C_D$. Comparing the results to those in Table 4.10, it can be seen that while a reduction in drag
4.3. DRAG SENSITIVITY

Figure 4.14: Optimum $C_D$ for varying tail radius ($r_t$) with fixed (blunt) nose geometry using the Blunted Wedge profile.

Figure 4.15: Optimum $C_D$ for varying tail length ($l_t$) with fixed (blunt) nose geometry using the Blunted Wedge profile.
CHAPTER 4. THE OPTIMIZATION OF SATELLITE’S AERODYNAMICS IN VLEO

(a) Original Cuboid ($r_t = 1, l_t = 0$)

(b) Partially Blunted Tail ($r_t = 0.5, l_t = 1$)

(c) Fully Pointed ($r_t = 0, l_t = 1$)

Figure 4.16: Normalised density field and surface force vectors around the tail geometry of the $0.5m \times 1.0m$ Blunted Wedge profile. The density field is normalised to the free-stream density (red dotted line). Pressure vectors (red) and force vectors (green) to scale and magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity.
4.3. DRAG SENSITIVITY

Table 4.11: Unconstrained optimized bodies for minimum $C_D$ for varying tail geometry and fixed (blunt) nose using the Blunted Wedge profile

<table>
<thead>
<tr>
<th>Profile Size $(height \times length)$</th>
<th>Cuboid $C_D$</th>
<th>Opti. Profile*</th>
<th>Opti. $C_D$</th>
<th>% Red. in Drag</th>
<th>Volume $m^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0.25m \times 0.5m$</td>
<td>2.435</td>
<td>▼</td>
<td>2.194</td>
<td>9.9</td>
<td>0.063</td>
</tr>
<tr>
<td>$0.25m \times 1.0m$</td>
<td>2.689</td>
<td>▼</td>
<td>2.237</td>
<td>16.8</td>
<td>0.125</td>
</tr>
<tr>
<td>$0.25m \times 1.5m$</td>
<td>2.933</td>
<td>▼</td>
<td>2.358</td>
<td>19.6</td>
<td>0.185</td>
</tr>
<tr>
<td>$0.50m \times 1.0m$</td>
<td>2.437</td>
<td>▼</td>
<td>2.192</td>
<td>10.0</td>
<td>0.250</td>
</tr>
<tr>
<td>$0.50m \times 1.5m$</td>
<td>2.567</td>
<td>▼</td>
<td>2.209</td>
<td>14.0</td>
<td>0.375</td>
</tr>
</tbody>
</table>

* direction of flight right-to-left with particle flow from the left

was achieved, it was about half the reduction achieved by the nose profile alone. This being the case, if the aim is to simply minimize the drag, a tapered nose would be favoured over a tapered tail as it provides a better reduction in $C_D$ for the same volume. This will be explored further in Sections 4.4 and 4.5. From Table 4.11 it can be seen that in the case of the fixed nose configuration, the optimum geometry is a tail with a sharp profile, as was seen in the case of the nose profiles. However, as discussed above and as was seen in Figure 4.14 at low tail radii, there is no significant change in the $C_D$ which could imply the existence of a line of local minima. While not shown here and in contrast with the profiles in Table 4.11, it was also observed that $r_t$ would fluctuate up and down as the fixed $l_t$ was varied. This appeared to be the case for all the aspect ratios tested which further supports the existence of a line of local minima.

The implication of this was that below a certain $r_t$ there is no further improvement in drag performance, a feature that is exemplified by the profiles with a 1:2 aspect ratio in Figure 4.14. Above 50%, the relationship appeared to be broadly exponential, which meant that for these bodies the majority of the $C_D$ reduction could be achieved with a relatively small reduction in $r_t$. For the other aspect ratios in Figure 4.14, the radius at which the plateauing occurred reduced as the aspect ratio of the body increased. This was possibly because these larger aspect ratios passed through a smaller range of tail angles, assuming $l_t$ was at its maximum. In cases where $l_t$ was limited, the flat region was much broader. This is a useful result since it implies that significant reductions in $C_D$ can be achieved with limited loss of internal volume. Being able to use larger tail radii has additional advantages, such as providing space for important equipment such as a propulsion system and a launch adaptor, which would typically be mounted to the rear surface of a satellite.

This raises the further question of how the reduction in the drag was achieved and why the $C_D$ appeared to plateau at lower tail radii. As discussed above, the particle velocities were not parallel and within the moving frame of the fluid have a random motion as described by the Maxwell-Boltzmann distribution [48]. These random motions had a component perpendicular to the flow...
which brought some of the particles into contact with the surface, causing an interaction and the transfer of energy. This in turn contributed towards the drag and resulted in the longer body’s experiencing high drag [3].

Another property of the gas was its high mean free path, which meant there was limited interaction between particles at this altitude and thus any object passing through the fluid at speed would leave a void behind it. The void would slowly refill, with the refill rate governed by the ratio between the body’s velocity and the perpendicular component of the random particle motions. The angle at which this occurs is the molecular Mach angle and can be calculated using the mean internal velocities from the Maxwell-Boltzmann distribution[92]. This is given by Equation 4.6 where the most probable thermal speed $c'$ is given by Equation 4.7 and where $v_\infty$ is the free stream speed, $k_B$ is the Boltzmann constant, $T$ is the temperature of the gas and $m$ is the particle mass of the gas species[92, 48].

$$
\theta_{refill} = \arctan \left( \frac{c'}{v_\infty} \right) 
$$

$$
v' = \sqrt{\frac{2k_B T}{m}} 
$$

By assuming these internal velocities act perpendicular to the flow, a very rough estimate of the refill angle for each particle species was derived, as presented in Table 4.12. Atomic Oxygen and Nitrogen comprise the highest percentage of particles in the fluid and thus contribute the most to the forces on the satellite. As can be seen from Table 4.12 they had a refill rate of between 6° – 8°. It would therefore be expected that if the tail angle was larger than 6° – 8°, there should only be a limited improvement in $C_D$. The data from Figure 4.14 was replotted against the internal
tail angle in Figure 4.17. From Figure 4.17, it can be seen that in general there was no further improvement in $C_D$ above about 7° of internal tail angle. There were a few anomalies near the right of the graph, which resulted from the shape tending towards a cuboid at high tail radii and thus being more susceptible to local minima. This trend was more obvious in bodies with lower aspect ratios, though high aspect ratio bodies appeared to follow the same trend. It should be noted that the plateau angles presented here were unique to the particular atmospheric conditions simulated. As described by the Maxwell-Boltzmann distribution, the random internal velocity of the particles is dependent on the temperature of the gas.

The rate of refill also has implications for other aspects of the satellite operation, for instance, the ingress of Atomic Oxygen to the interior of the satellite especially around sensitive equipment such as payloads. It was therefore of interest to properly describe the refill rate and explore its impact on other aspects of the satellite design. This aspect will be explored in more detail in Chapter 5.

Table 4.12: Approximate refill angles for gas species during the atmospheric condition described in Table 4.1

<table>
<thead>
<tr>
<th>Gas Species</th>
<th>Number Density [1/m]</th>
<th>Refill Angle [°]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atomic Oxygen ($O_1$)</td>
<td>$6.08 \times 10^{15}$</td>
<td>7.7</td>
</tr>
<tr>
<td>Nitrogen ($N_2$)</td>
<td>$4.31 \times 10^{15}$</td>
<td>5.8</td>
</tr>
<tr>
<td>Oxygen ($O_2$)</td>
<td>$3.21 \times 10^{14}$</td>
<td>5.5</td>
</tr>
<tr>
<td>Helium ($He$)</td>
<td>$6.33 \times 10^{13}$</td>
<td>15.1</td>
</tr>
<tr>
<td>Atomic Nitrogen ($N_1$)</td>
<td>$3.01 \times 10^{13}$</td>
<td>8.2</td>
</tr>
<tr>
<td>Argon ($Ar$)</td>
<td>$6.97 \times 10^{12}$</td>
<td>4.9</td>
</tr>
<tr>
<td>Hydrogen ($H$)</td>
<td>$1.43 \times 10^{12}$</td>
<td>28.3</td>
</tr>
</tbody>
</table>
4.3.4 Results & Discussion: Elliptical Profile

This section will describe and discuss how the $C_D$ changes as a result of the elements of the Elliptical profile, that being the length of the rounded nose and tail ($l_n$ & $l_t$ respectively). Figure 4.18 shows the variation in the $C_D$ with $l_n$ and with the tail held in a blunt configuration ($l_t = 0$). The minimum drag profiles with a fixed blunt tail are shown in Table 4.13 for the 5 body sizes, with their respective $C_D$ and the $C_D$ of their corresponding cuboid. Similarly, Figure 4.19 shows the variation in the $C_D$ with $l_t$ and with the nose held in a blunt configuration ($l_n = 0$). The minimum drag profiles with fixed blunt noses are shown in Table 4.14 for the 5 body sizes, also with their respective $C_D$.

The simulations performed for this section were carried out using DSMC and assumed a circular orbit of 200km altitude and energy accommodation coefficient $\alpha_T = 0.95$. Figure 4.20 shows the normalised density field as well as the surface pressure and force vectors around selected configurations of the Elliptical profile on a $0.5m \times 1.0m$ body. The density field is normalised to the density of the far-field as given in Table 4.1 (the dotted red line indicates regions of the density field that are equal to the far-field density). The surface pressure and force vectors are to scale but have been magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity. The examples presented in Figure 4.20 are the original cuboid ($l_n = 0, l_t = 0$) (Figure 4.20a), an Elliptical nose profile ($l_n = 1, l_t = 0$) (Figure 4.20b) and an Elliptical tail profile ($l_n = 0, l_t = 1$) (Figure 4.16c).

As with the Blunted Wedge, the lowest drag was achieved when the $l_n$ was at its maximum, in this case the full length of the body. However, as might be expected, the reduction in $C_D$ was not nearly as high. For instance, the 1:2 profiles achieved a minimum $C_D$ of 1.93 using just the nose of the Blunted Wedge in a sharp configuration, while the nose of the Elliptical profile only achieved a minimum of 2.04. The primary reason for this is the rounding of the nose tip. The rounded nose acts to blunt the profile, increasing the frontal area that was fully reflecting the flow back on itself, in a similar manner to the nose radius ($r_n$) of the Blunted Wedge. In fact, it can be inferred from Figures 4.12 & 4.18 that the Elliptical profile is approximately equivalent to a Blunted Wedge when $r_n = 30\%$. This once again raises questions whether in this case, the Elliptical Nose can provide any benefit. Like the rounded nose, the primary benefit is a potential improvement of the internal volume, however, this will be discussed in further detail in Sections 4.4 and 4.5.
4.3. DRAG SENSITIVITY

Figure 4.18: Optimum $C_D$ for varying nose length $(l_n)$ with fixed (blunt) tail geometry using the Elliptical profile

Table 4.13: Unconstrained optimized bodies for minimum $C_D$ for varying nose geometry and fixed (blunt) tail geometry using the Elliptical profile

<table>
<thead>
<tr>
<th>Profile Size (height x length)</th>
<th>Cuboid $C_D$</th>
<th>Opti. Profile*</th>
<th>Opti. $C_D$</th>
<th>% Red. in Drag</th>
<th>Volume $m^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0.25m \times 0.5m$</td>
<td>2.433</td>
<td></td>
<td>2.042</td>
<td>16.1</td>
<td>0.098</td>
</tr>
<tr>
<td>$0.25m \times 1.0m$</td>
<td>2.684</td>
<td></td>
<td>2.124</td>
<td>20.8</td>
<td>0.196</td>
</tr>
<tr>
<td>$0.25m \times 1.5m$</td>
<td>2.937</td>
<td></td>
<td>2.270</td>
<td>22.7</td>
<td>0.293</td>
</tr>
<tr>
<td>$0.50m \times 1.0m$</td>
<td>2.437</td>
<td></td>
<td>2.044</td>
<td>16.1</td>
<td>0.391</td>
</tr>
<tr>
<td>$0.50m \times 1.5m$</td>
<td>2.566</td>
<td></td>
<td>2.076</td>
<td>19.1</td>
<td>0.586</td>
</tr>
</tbody>
</table>

* direction of flight right-to-left with particle flow from the left
CHAPTER 4. THE OPTIMIZATION OF SATELLITE’S AERODYNAMICS IN VLEO

Figure 4.19: Optimum $C_D$ for varying tail length ($l_t$) with fixed (blunt) nose geometry using the Elliptical profile

Table 4.14: Unconstrained optimized bodies for minimum $C_D$ for varying tail geometry and fixed (blunt) nose geometry using the Elliptical profile

<table>
<thead>
<tr>
<th>Profile Size (height x length)</th>
<th>Cuboid $C_D$</th>
<th>Opti. Profile*</th>
<th>Opti. $C_D$</th>
<th>% Red. in Drag</th>
<th>Volume $m^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25m x 0.5m</td>
<td>2.434</td>
<td>-</td>
<td>2.235</td>
<td>8.2</td>
<td>0.0977</td>
</tr>
<tr>
<td>0.25m x 1.0m</td>
<td>2.684</td>
<td>-</td>
<td>2.358</td>
<td>12.1</td>
<td>0.195</td>
</tr>
<tr>
<td>0.25m x 1.5m</td>
<td>2.938</td>
<td>-</td>
<td>2.524</td>
<td>14.1</td>
<td>0.293</td>
</tr>
<tr>
<td>0.50m x 1.0m</td>
<td>2.437</td>
<td>-</td>
<td>2.236</td>
<td>8.3</td>
<td>0.391</td>
</tr>
<tr>
<td>0.50m x 1.5m</td>
<td>2.566</td>
<td>-</td>
<td>2.294</td>
<td>10.6</td>
<td>0.586</td>
</tr>
</tbody>
</table>

* direction of flight right-to-left with particle flow from the left
4.3. DRAG SENSITIVITY

Figure 4.20: Normalised density field and surface force vectors around the nose geometry of the 0.5m × 1.0m Elliptical profile. The density field is normalised to the free-stream density (red dotted line). Pressure vectors (red) and force vectors (green) to scale and magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity.
CHAPTER 4. THE OPTIMIZATION OF SATELLITE’S AERODYNAMICS IN VLEO

4.3.5 Results & Discussion: Double-Conic & Rounded-Conic Nose

This section will describe and discuss the sensitivity of the $C_D$ to the elements of the Double-Conic and Rounded-Conic nose profiles ($r_n$). In many respects, these profiles are quite similar, as they both allow the blunting of the front portion of the shape, and thus there are some similar parameters. Figure 4.21 shows the sensitivity of the $C_D$ to the total length of the nose profiles ($l_n$), Figure 4.21a for the Double-Conic profile and Figure 4.21b for the Rounded-Conic nose profiles. Similarly, Figure 4.21 shows the sensitivity of the $C_D$ to the radius of the nose ($r_n$), Figure 4.21a for the Double-Conic profile and Figure 4.21a for the Rounded-Conic nose profiles. The Double-Conic profile has a third parameter that controls the length of the front cone ($l_f$). The sensitivity of $C_D$ to this parameter is shown in Figure 4.23. The minimum drag configurations for the Double-Conic and Rounded-Conic nose profiles are shown in Tables 4.15 & 4.16 for the 5 body sizes with their respective $C_D$ and the $C_D$ of their corresponding cuboids.

The simulations performed for this section were carried out using DSMC and assumed a circular orbit of 200km altitude and an energy accommodation $\alpha_T = 0.95$. Figure 4.24 shows the normalised density field as well as the surface pressure and force vectors around examples of the Double-Conic & Rounded-Conic nose profiles on a 0.5m x 1.0m body. The density field is normalised to the density of the far-field as given in Table 4.1 (the dotted red line indicates regions of the density field that are equal to the far-field density). The surface pressure and force vectors are to scale but have been magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity. The examples presented in Figure 4.24 are the original cuboid ($r_n = 1, l_f = 0, l_n = 0$) (Figure 4.24a), a Double-Conic nose profile ($r_n = 0.5, l_f = 0.25, l_n = 1$) (Figure 4.24b) and Rounded-Conic nose profile ($r_n = 0.5, l_f = N/A, l_n = 1$) (Figure 4.16c).

From the configurations presented in Tables 4.15 & 4.16 it is apparent that from a purely drag perspective, there is no real benefit to either of these profiles at 200km altitude. Indeed on inspection of the graphs in Figure 4.21a, it can be seen that the results that they present are very similar to the result seen in Figure 4.11 which implies that the length of the tapered profile is the dominant factor in reducing the drag. Additionally, Figure 4.23 shows that there is no change in $C_D$ as the $l_f$ is varied. This is because the minimum drag is achieved when the surfaces of the front and main nose are coplanar, so the optimizer selects the narrowest $r_n$ possible, noting that the profile is prevented from becoming concave. Similarly, when $r_n$ is fixed, the optimizer should choose a $l_f$ that keeps the two surfaces coplanar. However, as can be seen from Figure 4.22a, there is a rise in $C_D$ above a radius of 25% of the body width. This is simply due to the constraints placed on the parameter that controls $l_f$, that this portion of the profile could not be longer than half the length of the body. This means that above a radius of 25% of the body’s width, the optimizer is not able to keep the surfaces coplanar and thus the larger front surface is turned back into the flow and the drag increases. So this would suggest that neither $r_n$ nor $l_f$ is effective at reducing the drag on the body for the Double-Conic profile.

By contrast, the Rounded-Conic nose profile appears to have a lot more variation. Comparing the results in Figure 4.22b with those of the Blunted Wedge in Figure 4.12, it can be seen that in general, the Rounded-Conic nose achieves a lower drag for a given $r_n$. This is because only the very
4.3. DRAG SENSITIVITY

(a) Double-Conic nose

(b) Rounded-Conic nose

Figure 4.21: Sensitivity of the $C_D$ to the total length of the nose ($l_n$)
CHAPTER 4. THE OPTIMIZATION OF SATELLITE'S AERODYNAMICS IN VLEO

Figure 4.22: Sensitivity of the $C_D$ to the front nose radius ($r_n$)
4.3. DRAG SENSITIVITY

front of the Rounded-Conic nose is pointed upstream and fully reflecting the flow back on itself. Since the rest of the profile is curved, the drag contribution slowly reduces until the surface is tangent to the tapered surface. This means the profile has a smaller equivalent blunt nose radius and therefore may perform better than the purely blunted wedge when the volume is considered. This will be explored in more detail in later sections.

In many respects, the results presented in this section are expected given prior observations. In Section 4.3.2, it was seen that the $C_D$ was minimized when either the tapered part of the satellite body was maximized or the bluntness of the nose was minimized. In this instance, these profiles attempted to temper the bluntness of the Wedge profile by either allowing the blunted nose section to taper itself or by simply rounding the nose entirely. While not ideal for reducing $C_D$, as was seen for the Rounded-Conic nose profile, these features may still provide advantages when also attempting to maximize the internal volume of the profile by reducing the effective $r_n$ of the profile. This will be explored in more detail in Sections 4.4 and 4.5.
### Table 4.15: Unconstrained optimized bodies for minimum $C_D$ for varying nose geometry using Double-Conic profile

<table>
<thead>
<tr>
<th>Profile Size ($height \times length$)</th>
<th>Cuboid $C_D$</th>
<th>Opti. Profile*</th>
<th>Opti. $C_D$</th>
<th>% Red. in Drag</th>
<th>Volume $m^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25m x 0.5m</td>
<td>2.389</td>
<td></td>
<td>2.042</td>
<td>14.5</td>
<td>0.063</td>
</tr>
<tr>
<td>0.25m x 1.0m</td>
<td>2.644</td>
<td></td>
<td>2.067</td>
<td>21.8</td>
<td>0.124</td>
</tr>
<tr>
<td>0.25m x 1.5m</td>
<td>2.914</td>
<td></td>
<td>2.184</td>
<td>25.1</td>
<td>0.187</td>
</tr>
<tr>
<td>0.50m x 1.0m</td>
<td>2.388</td>
<td></td>
<td>2.042</td>
<td>14.5</td>
<td>0.253</td>
</tr>
<tr>
<td>0.50m x 1.5m</td>
<td>2.519</td>
<td></td>
<td>2.043</td>
<td>19.0</td>
<td>0.375</td>
</tr>
</tbody>
</table>

*direction of flight right-to-left with particle flow from the left

### Table 4.16: Unconstrained optimized bodies for minimum $C_D$ for varying nose geometry using Rounded-Conic nose

<table>
<thead>
<tr>
<th>Profile Size ($height \times length$)</th>
<th>Cuboid $C_D$</th>
<th>Opti. Profile*</th>
<th>Opti. $C_D$</th>
<th>% Red. in Drag</th>
<th>Volume $m^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25m x 0.5m</td>
<td>2.391</td>
<td></td>
<td>2.050</td>
<td>14.3</td>
<td>0.067</td>
</tr>
<tr>
<td>0.25m x 1.0m</td>
<td>2.653</td>
<td></td>
<td>2.073</td>
<td>21.9</td>
<td>0.128</td>
</tr>
<tr>
<td>0.25m x 1.5m</td>
<td>2.915</td>
<td></td>
<td>2.197</td>
<td>24.6</td>
<td>0.193</td>
</tr>
<tr>
<td>0.50m x 1.0m</td>
<td>2.391</td>
<td></td>
<td>2.045</td>
<td>14.5</td>
<td>0.253</td>
</tr>
<tr>
<td>0.50m x 1.5m</td>
<td>2.522</td>
<td></td>
<td>2.056</td>
<td>18.5</td>
<td>0.412</td>
</tr>
</tbody>
</table>

*direction of flight right-to-left with particle flow from the left
4.3. DRAG SENSITIVITY

(a) Original Cuboid \( (r_n = 1, l_f = 0, l_n = 0) \)

(b) Double-Conic nose \( (r_n = 0.5, l_f = 0.25, l_n = 1) \)

(c) Rounded-Conic nose \( (r_n = 0.5, l_f = N/A, l_n = 1) \)

Figure 4.24: Normalised density field and surface force vectors around the nose geometry of the 0.5m × 1.0m Double-Conic & Rounded-Conic nose. The density field is normalised to the free-stream density (red dotted line). Pressure vectors (red) and force vectors (green) to scale and magnified by \( 1 \times 10^4 \) and \( 1 \times 10^5 \) respectively for clarity.
4.4 Fixed Body Dimensions

4.4.1 Introduction

In the previous sections, the sensitivity of the aerodynamic drag to each individual element of the profiles was explored. In general, it was demonstrated that to purely minimize the drag, the most appropriate shape would be a simple conical profile, the additional geometric features explored provided no additional benefit to this scenario. While drag is an important aspect of the design, as was also seen in the last section, its minimization comes at the expense of internal volume. In this section, optimum geometries will be identified by trading off drag against the internal volume of the profiles applying multi-objective optimization. In the first instance, this is done for the fixed external dimensions as described in Section 4.2.5.4 for all four profiles.

4.4.2 Results & Discussion

The graphs in Figures 4.25, 4.27, 4.29 & 4.31 show the Pareto-optimal fronts for the four profiles Blunted Wedge, Elliptical, Double-Conic and Rounded-Conic respectively. These were obtained by varying the relative weight of the two objectives using the $\alpha$ variable in Equation (4.3). These are the lines of optimal solutions whose objective functions cannot be improved upon without detriment to at least one of the objectives. In each figure, a separate Pareto-optimal front is shown for each of the five fixed external dimensions described in Section 4.2.3. Examples of the shapes achieved across the fronts for the bodies with a 1:2 (left) and 1:6 (right) aspect ratio are provided in Tables 4.17, 4.18, 4.19 & 4.20 for the four profiles Blunted Wedge, Elliptical, Double-Conic and Rounded-Conic respectively. The top profile in each table represents bodies optimized purely for maximum volume ($\alpha = 0$), while the lowest bodies were optimized purely for minimizing drag ($\alpha = 1$). It should also be noted that Figure 4.27 only shows data up to a maximum reduction of 22% for all body sizes. This is simply because this is the maximum reduction possible given the morphology of the Elliptical profile. By contrast, the other three profiles, the Blunted Wedge, Double-Conic and Rounded-Conic, can form into a simple cone given the right parameters which in all cases contains half the available volume of the equivalent cuboid.

Figures 4.26, 4.28, 4.30 & 4.32 show the normalised density field, as well as the surface pressure and force vectors around example Pareto-optimal profiles for the Blunted Wedge, Elliptical, Double-Conic and Rounded-Conic respectively on a $0.5m \times 1.0m$ body. The density field is normalised to the density of the far-field as given in Table 4.1 (the dotted red line indicates regions of the density field that are equal to the far-field density). The surface pressure and force vectors are to scale but have been magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity.

As can be seen from Figures 4.25 & 4.31, the resultant Pareto-optimal fronts were broadly exponential for all four profiles. In general, this meant that, as the volume is reduced, so do potential improvements in the drag. This is exemplified by the 1:2 Blunted Wedge which, as can be seen in Table 4.17 and Figure 4.25, achieves a 13.0% reduction in the drag for a 30% reduction in the internal volume while a 40% or 50% reduction in volume still only achieves a 13.3% in the drag. The implication is that for a 30% reduction in volume the Blunted Wedge profile has achieved over 90% of its overall achievable drag reduction and indeed most of the improvement in the drag is...
achieved by reducing the volume by just 10%. Similar behaviour is observed for the larger aspect ratios, though the point of diminishing returns is at a lower volume relative to the equivalent cuboid.

It can also be seen that the ‘Nose-only’ profiles (the Double-Conic (Figure 4.29) and Rounded-Conic (Figure 4.31) profiles) do not achieve the same overall reduction in the drag as the Blunted Wedge (Figure 4.25). This is particularly true for the larger aspect ratio bodies, such as the 1:6 bodies. In this instance, the Blunted Wedge achieves a maximum drag reduction of 30% (2.05) compared to 25% (2.19) for the two ‘Nose-only’ profiles. This is due to the incorporation of the tapered tail in the profile which, as was discussed in Section 4.3.3, effectively reduces the length of the profile by turning the rear surfaces away from the flow. Similar results are true for 1:6 Elliptical profile which similarly achieves a lower $C_D$ for a given volume than either of the ‘Nose-only’ profiles, employing its tail section to do this (as can be seen in Table 4.18).

Interestingly from Table 4.18, it can be seen that for a given volume the Elliptical profile appears to achieve a lower drag than the Blunted Wedge profile. This was slightly unexpected since the rounded profiles were originally included just for interest and completeness. There are three reasons this profile might perform better than the Blunted Wedge for a given volume. Firstly, unlike the Blunted Wedge, only the tip of the nose is fully facing the flow, so there is a gradual reduction in the drag contribution the further from the mid-line the point on the surface is. This continues until the nose surface is tangential to the side panels, which highlights the second reason: these surfaces near the edge contribute very little to the overall drag due to their low internal cone angle. Finally, since the surface is curved and bulbous, it can accommodate significantly more than a comparable flat-plated Blunted Wedge. For instance, to achieve roughly the same $C_D$ as the 1:2 Elliptical profile with a volume of 0.4 $m^3$ the 1:2 Blunted Wedge has to reduce to at least 0.3 $m^3$ as can be seen from Tables 4.17 & 4.18.

Similar observations can be made of the Double-Conic profile in Table 4.29. To strike a balance between the drag and the volume, and in contrast to the Blunted Wedge, the front nose of the Double-Conic profile is also tapered but with a higher half angle. This means that, for a given radius, the front surfaces do not experience as great a contribution to the overall drag as the flat noses of the Blunted Wedge. In many respects, the Double-Conic profile is simply a lower resolution Elliptical nose, which is why there are some similarities in the $C_D$ for the lower aspect ratio shapes of these profiles where the tail section is not utilized.

The four profiles, while distinct from each other, have fixed topologies. The parameters can stretch and alter the placement of the vertices, but fundamentally a Blunted Wedge profile could not reproduce the profiles produced by the Elliptical profile. It is, however, interesting to consider what optimized geometries could be achieved if the vertices were controlled directly. As it happens, the vertices that locate the transition from the front nose to the rear nose is controlled by two parameters $l_f$ and $r_n$, which control the x and y location respectively. Based on these results, it is likely that a free-form geometry would attempt to minimize the frontal area pointed directly by chamfering the edges or smoothly tapering the surface back. This would be an interesting area for future work and a natural progression from work presented here.
CHAPTER 4. THE OPTIMIZATION OF SATELLITE’S AERODYNAMICS IN VLEO

Figure 4.25: Graph showing the Pareto-optimal fronts for the Blunted Wedge profile when optimizing for minimum drag and maximum volume

Table 4.17: Example optimal Blunted Wedge profiles at varying points along the Pareto-optimal fronts for 1:2 profiles (left) and 1:6 profiles (right)

<table>
<thead>
<tr>
<th>1:2 Shape*</th>
<th>Volume $[m^3]$</th>
<th>$C_D$ [-]</th>
<th>[％]</th>
<th>1:6 Shape*</th>
<th>Volume $[m^3]$</th>
<th>$C_D$ [-]</th>
<th>[％]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.50</td>
<td>2.392</td>
<td>0.00</td>
<td>0.38</td>
<td>2.915</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.45</td>
<td>2.144</td>
<td>10.36</td>
<td>0.34</td>
<td>2.486</td>
<td>14.75</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.40</td>
<td>2.092</td>
<td>12.53</td>
<td>0.30</td>
<td>2.309</td>
<td>20.81</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.35</td>
<td>2.073</td>
<td>13.33</td>
<td>0.26</td>
<td>2.174</td>
<td>24.42</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.30</td>
<td>2.066</td>
<td>13.62</td>
<td>0.23</td>
<td>2.088</td>
<td>27.38</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>2.066</td>
<td>13.62</td>
<td>0.19</td>
<td>2.046</td>
<td>29.82</td>
<td></td>
</tr>
</tbody>
</table>

* direction of flight right-to-left with particle flow from the left
4.4. FIXED BODY DIMENSIONS

(a) Volume Reduction = 0%

(b) Volume Reduction = 10%

(c) Volume Reduction = 20%
Figure 4.26: Normalised density field and surface force vectors around the nose geometry of the 0.5m × 1.0m Blunted Wedge profile. The density field is normalised to the free-stream density (red dotted line). Pressure vectors (red) and force vectors (green) to scale and magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity.
Figure 4.27: Graph showing the Pareto-optimal fronts for the Elliptical profile when optimizing for minimum drag and maximum volume

Table 4.18: Example optimal Elliptical profiles at varying points along the Pareto-optimal fronts for 1:2 profiles (left) and 1:6 profiles (right)

<table>
<thead>
<tr>
<th>1:2 Shape*</th>
<th>Volume [m$^3$]</th>
<th>$C_D$ [-]</th>
<th>1:6 Shape*</th>
<th>Volume [m$^3$]</th>
<th>$C_D$ [-]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(%)</td>
<td></td>
<td></td>
<td>(%)</td>
<td></td>
</tr>
<tr>
<td>0.50</td>
<td>0</td>
<td>2.401</td>
<td>0.38</td>
<td>0</td>
<td>2.921</td>
</tr>
<tr>
<td>0.47</td>
<td>5</td>
<td>2.300</td>
<td>0.36</td>
<td>5</td>
<td>2.621</td>
</tr>
<tr>
<td>0.45</td>
<td>10</td>
<td>2.198</td>
<td>0.34</td>
<td>10</td>
<td>2.454</td>
</tr>
<tr>
<td>0.42</td>
<td>15</td>
<td>2.123</td>
<td>0.32</td>
<td>15</td>
<td>2.319</td>
</tr>
<tr>
<td>0.40</td>
<td>20</td>
<td>2.062</td>
<td>0.30</td>
<td>20</td>
<td>2.207</td>
</tr>
</tbody>
</table>

* direction of flight right-to-left with particle flow from the left
Figure 4.28: Normalised density field and surface force vectors around the nose geometry of the 0.5m × 1.0m Elliptical profile. The density field is normalised to the free-stream density (red dotted line). Pressure vectors (red) and force vectors (green) to scale and magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity.
Figure 4.29: Graph showing the Pareto-optimal fronts for the Double-Conic profile when optimizing for minimum drag and maximum volume

Table 4.19: Example optimal Double-Conic profiles at varying points along the Pareto-optimal fronts for 1:2 profiles (left) and 1:6 profiles (right)

<table>
<thead>
<tr>
<th>1:2 Shape*</th>
<th>Volume</th>
<th>$C_D$</th>
<th>1:6 Shape*</th>
<th>Volume</th>
<th>$C_D$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[m$^3$]</td>
<td>[%]</td>
<td>[-]</td>
<td>[m$^3$]</td>
<td>[%]</td>
</tr>
<tr>
<td>0.50</td>
<td>0</td>
<td>2.392</td>
<td>0.37</td>
<td>1.36</td>
<td>0.00</td>
</tr>
<tr>
<td>0.45</td>
<td>10</td>
<td>2.202</td>
<td>0.34</td>
<td>1.15</td>
<td>11.20</td>
</tr>
<tr>
<td>0.40</td>
<td>20</td>
<td>2.125</td>
<td>0.30</td>
<td>1.08</td>
<td>16.38</td>
</tr>
<tr>
<td>0.35</td>
<td>30</td>
<td>2.077</td>
<td>0.26</td>
<td>0.94</td>
<td>20.45</td>
</tr>
<tr>
<td>0.30</td>
<td>40</td>
<td>2.094</td>
<td>0.22</td>
<td>0.91</td>
<td>23.36</td>
</tr>
<tr>
<td>0.25</td>
<td>50</td>
<td>2.042</td>
<td>0.19</td>
<td>0.88</td>
<td>25.05</td>
</tr>
</tbody>
</table>

* direction of flight right-to-left with particle flow from the left
CHAPTER 4. THE OPTIMIZATION OF SATELLITE’S AERODYNAMICS IN VLEO

(a) Volume Reduction = 0%

(b) Volume Reduction = 10%

(c) Volume Reduction = 20%
4.4. FIXED BODY DIMENSIONS

(d) Volume Reduction = 30%

(e) Volume Reduction = 40%

(f) Volume Reduction = 50%

Figure 4.30: Normalised density field and surface force vectors around the nose geometry of the 0.5m × 1.0m Double-Conic profile. The density field is normalised to the free-stream density (red dotted line). Pressure vectors (red) and force vectors (green) to scale and magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity.
Figure 4.31: Graph showing the Pareto-optimal fronts for the Rounded-Conic profile when optimizing for minimum drag and maximum volume

Table 4.20: Example optimal Rounded-Conic profiles at varying points along the Pareto-optimal fronts for 1:2 profiles (left) and 1:6 profiles (right)

<table>
<thead>
<tr>
<th>1:2 Shape*</th>
<th>Volume [m³]</th>
<th>C_D [-]</th>
<th>1:6 Shape*</th>
<th>Volume [m³]</th>
<th>C_D [-]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[%]</td>
<td></td>
<td></td>
<td>[%]</td>
<td></td>
</tr>
<tr>
<td>0.50</td>
<td>0.00</td>
<td>2.391</td>
<td>0.37</td>
<td>0.00</td>
<td>2.915</td>
</tr>
<tr>
<td>0.45</td>
<td>10.00</td>
<td>2.210</td>
<td>0.34</td>
<td>10.00</td>
<td>2.577</td>
</tr>
<tr>
<td>0.40</td>
<td>20.00</td>
<td>2.121</td>
<td>0.30</td>
<td>20.00</td>
<td>2.446</td>
</tr>
<tr>
<td>0.35</td>
<td>30.00</td>
<td>2.077</td>
<td>0.26</td>
<td>30.00</td>
<td>2.341</td>
</tr>
<tr>
<td>0.30</td>
<td>40.00</td>
<td>2.056</td>
<td>0.22</td>
<td>40.00</td>
<td>2.261</td>
</tr>
<tr>
<td>0.25</td>
<td>50.00</td>
<td>2.043</td>
<td>0.19</td>
<td>50.00</td>
<td>2.186</td>
</tr>
</tbody>
</table>

*direction of flight right-to-left with particle flow from the left
4.4. FIXED BODY DIMENSIONS

(a) Volume Reduction = 0%

(b) Volume Reduction = 10%

(c) Volume Reduction = 20%
Figure 4.32: Normalised density field and surface force vectors around the nose geometry of the 0.5m × 1.0m Rounded-Conic profile. The density field is normalised to the free-stream density (red dotted line). Pressure vectors (red) and force vectors (green) to scale and magnified by $1 \times 10^4$ and $1 \times 10^5$ respectively for clarity.
So far, the emphasis in this discussion has been on the comparative merits of each profile for given reductions in the internal volume. This is because it has been assumed that the external dimensions are fixed and that the drag can only be reduced by trading off against the internal volume. This presents some difficulties since reducing the volume directly impacts the size of the subsystems that the aeroshell can accommodate. This raises the question, what loss in internal volume would actually be acceptable? Fortunately, one of the key drivers of this work is to reduce the fuel mass required for a given mission profile.

In general, the fuel mass and its volume is proportional to the drag the platform experiences, so it can be assumed that a reduction in the drag will result in an equally proportioned reduction in required fuel system volume (see Equation 2.5). This is an oversimplification since there will be elements that will not scale with the fuel and individual configurations may differ in scaling, but this will suffice for the scope of this work. Based on this assumption, for the profiles to be viable, the reduction in the required fuel system volume (or the drag) should be equal to or less than the loss of internal volume due to the aeroshell. This boundary is represented by the diagonal dashed line in Figures 4.25-4.31, which assumes a 1:1 scaling. Those points above and to the left have a net penalty, with fuel volume potentially using space for the payload and other subsystems. In general, there are two points at which the Pareto-optimal fronts cross this line. The first is when there is zero reduction in both the volume and the drag, which in all cases is the equivalent cuboid, so not of immediate interest here. The second points are shown in Table 4.21 for all 4 profiles and all 5 aspect ratios. In some cases, such as the 1:2 Rounded-Conic nose profiles, the drag is never sufficiently reduced to provide a net benefit of using the profile. In other cases, such as the 1:4 and 1:6 Elliptical profile, the volume reduction limit of 22% is reached before a second crossing is made, the values in brackets indicate the maximum possible drag reduction in this case.

Table 4.21: Maximum % reduction in $C_D$ achievable with no net penalty assuming the fuel system volume scales with the drag value

<table>
<thead>
<tr>
<th>Profile Size (height x length)</th>
<th>Blunted Wedge</th>
<th>Elliptical</th>
<th>Double-Conic</th>
<th>Rounded-Conic</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25m x 0.5m</td>
<td>10.6</td>
<td>9.3</td>
<td>6.2</td>
<td>--</td>
</tr>
<tr>
<td>0.25m x 1.0m</td>
<td>17.5</td>
<td>(22.3)</td>
<td>10.1</td>
<td>10.4</td>
</tr>
<tr>
<td>0.25m x 1.5m</td>
<td>21.6</td>
<td>(25.9)</td>
<td>12.8</td>
<td>13.5</td>
</tr>
<tr>
<td>0.50m x 1.0m</td>
<td>10.8</td>
<td>9.3</td>
<td>6.2</td>
<td>--</td>
</tr>
<tr>
<td>0.50m x 1.5m</td>
<td>14.4</td>
<td>17.1</td>
<td>8.4</td>
<td>8.1</td>
</tr>
</tbody>
</table>

From the data presented in Table 4.21 it can be seen that, in general, an optimized aeroshell provides more benefit for longer bodies when the external dimensions are fixed, as they can achieve a much greater reduction in drag before making a net loss as a result of lost volume. Once again it can also be seen that the Blunted Wedge and Elliptical profiles generally perform better than the ‘Nose-only’ profiles. This is primarily due to the inclusion of the tail section since it effectively reduces the length of the body presented to the flow while maintaining a greater volume. Unsurprisingly, this is most effective for the longer bodies which can achieve a reduction in the drag that is greater than 20%. It can also be seen that the Elliptical profile performs best overall in this case, for the reasons described above. While the Elliptical profile may be the more efficient
profile, there is still the question of manufacture and integration, since flat structural panels are easier to manufacture than curved panels and other systems can be more readily mounted to them. Therefore it would be up to the system engineer to decide whether a 5% reduction in drag (over the Blunted Wedge profile) was worth the additional manufacturing hurdles that a set of curved surfaces would create.

4.4.3 Summary

In this section, optimum geometries were identified by trading off the drag against the internal volume of the profiles by using multi-objective optimization for a set of fixed external dimensions (described in Section 4.2.3). It was seen that the Elliptical profile generally performed the best of the 4 profiles, particularly for long higher aspect ratio profiles. This was especially the case when considering the acceptable losses, where it was seen that for the 1:6 profiles a reduction of 26% was achievable with a reduction in the volume of just 22%. This compares to the 1:6 Blunted Wedge profile that achieved a reduction in both the drag and volume of 22%. It was also generally observed that the ‘Nose-only’ profiles did not perform as well as the Blunted Wedge or Elliptical profiles which both featured a tail section. This was the case even when the optimizer was solely optimizing for drag, and the key difference was that the tail section effectively shortened the profiles by turning the rear surfaces away from the flow.
4.5 Fixed Volume

4.5.1 Introduction

In the previous section, the optimum geometries for fixed external dimensions were considered. While the external dimensions can often be a constraining factor on the geometry of the satellite, it is more common to prescribe a desired internal volume for the satellite body. In this section, results will be presented for a range of required internal volumes for each profile as described in Section 4.2.5.5. Left unhindered and given the results seen so far, the optimizer would likely opt for the thinnest profile available for the data sets. This may not be desirable, so to simulate the dimensional needs of internal subsystems, fixed body lengths and width will be set separately. The simulations performed for this section were carried out using DSMC and assumed a circular orbit of 200km altitude and $a = 0.95$.

4.5.2 Fixed volume with fixed body length

The graphs in Figure 4.33 show the variation in the drag force with internal volume for bodies of constant length, between 0.5m and 1.5m. As previously, this is broken down into the four profiles: Blunted Wedge (Figure 4.33a), Elliptical (Figure 4.33b), Double-Conic (Figure 4.33c) and Rounded-Conic (Figure 4.33d). In each case, the coloured dashed line represents the drag force of the equivalent cuboid profile, i.e. the drag a cuboid profile of similar volume would experience when equally constrained. Also shown in the figures is the approximate extent of the data sets given by the black dotted line. The improvement in the drag force as compared to the equivalent cuboid profile is provided in Figure 4.34 for each profile.

While the trend of the data is quite clear, there are notable deviations in the plots in Figure 4.33, which are even more apparent in Figure 4.34. These are likely the result of poor recovery of the original function by the surrogate model, for instance by introducing additional local minima. This was a concern for data sets with a larger number of parameters since the number of samples to effectively sample the design space is roughly of $O(a^m)$ where $m$ is the number of parameters. While the total number of samples will still be much smaller than an equivalent gridded sampling method, the number of samples needed can still become very large (see Section 3.3.2). So a data set that is smaller than ideal has to be accepted, but this runs the risk that there will be regions that have a lower sampling density and are thus not as well covered. In the previous sections, this could be mitigated by carefully choosing the body dimension or reinforcing the surrogate model using a data set with a reduced number of parameters.

As can be seen from Figure 4.34, when the length of the satellite body is fixed, adopting one of these profiles would only provide a marginal improvement on the cuboid profile. The rounded nose profile for instance (Figure 4.34d) showed no meaningful improvement in this scenario. The reason for this is that the drag on a satellite is normally dictated by the frontal projected area, so an increase in this area would normally result in a proportional increase in the drag on the body. However, as can be seen in Table 4.22, to achieve the desired volume and incorporate the profile, the width of the body must also grow. This means that the problem becomes a trade-off between the drag-reducing effects of the profiles seen in Section 4.3 and the drag-increasing effects of the
Figure 4.33: Variation of atmospheric drag with volume when the length of the body is fixed, comparing Cuboid profile (dashed line) to a Blunted Wedge (a), Elliptical profile (b), Double-Conic (c) and Rounded-Conic (d)
Figure 4.34: Improvement in atmospheric drag compared to a Cuboid profile when the length of the body is fixed
Table 4.22: Comparison of the atmospheric drag of the Blunted Wedge and Elliptical profiles with the simple Cuboid profile given the desired volume and when the length of the profile is fixed at 1m

<table>
<thead>
<tr>
<th>req. vol. $[m^3]$</th>
<th>Cuboid</th>
<th>Blunted Wedge</th>
<th>Elliptical</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>width $[m]$</td>
<td>drag $[mN]$</td>
<td>width $[m]$</td>
</tr>
<tr>
<td>0.15</td>
<td>0.15</td>
<td>5.15</td>
<td>0.20</td>
</tr>
<tr>
<td>0.20</td>
<td>0.20</td>
<td>6.42</td>
<td>0.22</td>
</tr>
<tr>
<td>0.25</td>
<td>0.25</td>
<td>7.65</td>
<td>0.26</td>
</tr>
<tr>
<td>0.30</td>
<td>0.30</td>
<td>8.88</td>
<td>0.32</td>
</tr>
<tr>
<td>0.35</td>
<td>0.35</td>
<td>10.11</td>
<td>0.37</td>
</tr>
<tr>
<td>0.40</td>
<td>0.40</td>
<td>11.33</td>
<td>0.42</td>
</tr>
<tr>
<td>0.45</td>
<td>0.45</td>
<td>12.56</td>
<td>0.47</td>
</tr>
<tr>
<td>0.50</td>
<td>0.50</td>
<td>13.80</td>
<td>0.52</td>
</tr>
</tbody>
</table>

1 direction of flight right-to-left with particle flow from the left
2 improvement relative to the equivalent Cuboid

body width. In this instance, the profiles are not sufficiently effective to offset the increased drag required to accommodate them. It should be noted that this is assuming $\alpha = 0.95$, so the profiles were experiencing mostly diffuse particle-surface interactions. The greater force experienced by the cuboid profile under specular conditions might increase the relative benefit of the conic profiles such as the Blunted Wedge. However, it is also known that as the surfaces degrade over time in the atomic oxygen-rich environment of VLEO, the surface interactions becomes more diffuse. So this discussion may ultimately be moot unless a material is identified that can maintain its specular behaviour for the duration of a mission (1 year or more).

Of the profiles, the Blunted Wedge performed the best across the range of required volumes, for instance, the 1m body achieved a maximum reduction of 6.87% but dropped as low as 2.81% for the longer bodies. To achieve this, the optimizer settled on a profile that was between 5 and 10% wider than the equivalent cuboid profile, with what amounted to chamfered corners, i.e. the nose and tail radius was large compared to the body. It is interesting to note that the pointed profiles, as seen in the previous section, were not effective here, as it would have required a much larger increase in the width of the body to accommodate the desired volume. Instead, the optimizer tapered the tail as well as the nose in almost all cases. As has previously been described, the tail profile reduces the effective length of the body, which for longer profiles can be an efficient way of reducing the drag without sacrificing large amounts of internal volume. However, as has been seen in the previous section, the tail profile is not normally as effective as the nose profile at reducing the drag for a given size of the body.
Table 4.23: Comparison of the atmospheric drag of the Double-Conic and Rounded-Conic profiles with the simple Cuboid profile given the desired volume and when the length of the profile is fixed at 1m

<table>
<thead>
<tr>
<th>req. vol. $[m^3]$</th>
<th>Cuboid</th>
<th>Double-Conic</th>
<th>Rounded-Conic</th>
</tr>
</thead>
<tbody>
<tr>
<td>width $[m]$</td>
<td>drag $[mN]$</td>
<td>width $[m]$</td>
<td>drag $[mN]$</td>
</tr>
<tr>
<td>0.15</td>
<td>0.15</td>
<td>5.15</td>
<td>0.20</td>
</tr>
<tr>
<td>0.20</td>
<td>0.20</td>
<td>6.42</td>
<td>0.22</td>
</tr>
<tr>
<td>0.25</td>
<td>0.25</td>
<td>7.65</td>
<td>0.27</td>
</tr>
<tr>
<td>0.30</td>
<td>0.30</td>
<td>8.88</td>
<td>0.31</td>
</tr>
<tr>
<td>0.35</td>
<td>0.35</td>
<td>10.11</td>
<td>0.36</td>
</tr>
<tr>
<td>0.40</td>
<td>0.40</td>
<td>11.33</td>
<td>0.40</td>
</tr>
<tr>
<td>0.45</td>
<td>0.45</td>
<td>12.56</td>
<td>0.45</td>
</tr>
<tr>
<td>0.50</td>
<td>0.50</td>
<td>13.80</td>
<td>0.50</td>
</tr>
</tbody>
</table>

$^1$ direction of flight right-to-left with particle flow from the left
$^2$ improvement relative to the equivalent Cuboid

The question therefore becomes, whether a 5% reduction in the drag force is worth the added structural and manufacturing complexity of incorporating the profile. While this would correspond with a roughly 5% reduction in the fuel required for drag compensation (assuming a fixed mission length), the added structural complexity might reduce the overall mass saving. Additionally, there is no guarantee that the internal space is usable, especially in the tapered regions. In short, the improvements in the drag that these profiles afford would need to be critically weighed up against the other system requirements on a case by case basis.

4.5.3 Fixed body width

The graphs in Figure 4.35 show the variation in the drag force with internal volume for bodies of constant width, between 0.2m and 0.7m. This is broken down into the four profiles: Blunted Wedge (Figure 4.35a), Elliptical profile (Figure 4.35b), Double-Conic (Figure 4.35c) and Rounded-Conic (Figure 4.35d). In each case, the coloured dashed line represents the drag force of the equivalent cuboid profile, i.e. the drag a cuboid profile of similar volume would experience when equally constrained. Also shown in the figures is the approximate extent of the data sets (given by the black dotted line). The improvement in the drag force, as compared to the equivalent cuboid profile is provided in Figure 4.36 for each profile.

As can be seen from Figure 4.36, a far greater reduction in the drag is achieved with fixed body width compared to when the length of the body is fixed, with improvements over the cuboid profile of as much as 15-20%. As previously discussed, the width of the body, or more specifically the
Figure 4.35: Variation of atmospheric drag with volume when the width of the body is fixed, comparing Cuboid profile (dashed line) to a Blunted Wedge (a), Elliptical profile (b), Double-Conic (c) and Rounded-Conic (d)
Figure 4.36: Improvement in atmospheric drag compared to a Cuboid profile when the width of the body is fixed
CHAPTER 4. THE OPTIMIZATION OF SATELLITE’S AERODYNAMICS IN VLEO

Table 4.24: Comparison of the atmospheric drag of the Blunted Wedge and Elliptical profiles with the simple Cuboid profile given the desired volume and when the width of the profile is fixed at 0.5m

<table>
<thead>
<tr>
<th>req. vol. [m^3]</th>
<th>Cuboid</th>
<th>Blunted Wedge</th>
<th>Elliptical</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>length [m]</td>
<td>drag [mN]</td>
<td>length [m]</td>
</tr>
<tr>
<td>0.05</td>
<td>0.10</td>
<td>12.45</td>
<td>0.20</td>
</tr>
<tr>
<td>0.10</td>
<td>0.20</td>
<td>12.60</td>
<td>0.40</td>
</tr>
<tr>
<td>0.15</td>
<td>0.30</td>
<td>12.74</td>
<td>0.60</td>
</tr>
<tr>
<td>0.20</td>
<td>0.40</td>
<td>12.89</td>
<td>0.80</td>
</tr>
<tr>
<td>0.25</td>
<td>0.50</td>
<td>13.03</td>
<td>1.00</td>
</tr>
<tr>
<td>0.30</td>
<td>0.60</td>
<td>13.19</td>
<td>1.22</td>
</tr>
<tr>
<td>0.35</td>
<td>0.70</td>
<td>13.34</td>
<td>1.41</td>
</tr>
<tr>
<td>0.40</td>
<td>0.80</td>
<td>13.50</td>
<td>1.49</td>
</tr>
<tr>
<td>0.45</td>
<td>0.90</td>
<td>13.65</td>
<td>1.50</td>
</tr>
<tr>
<td>0.50</td>
<td>1.00</td>
<td>13.80</td>
<td>1.50</td>
</tr>
</tbody>
</table>

1 direction of flight right-to-left with particle flow from the left
2 improvement relative to the equivalent Cuboid

Frontal projected area is the primary indicator of the force that that body will experience within a rarefied gas. However, the drag also increases in proportion to the length of the body, though at a much-reduced rate, which can be seen in the cuboid profiles in Figure 4.35. All four profiles appear to counteract the effect of the length entirely, despite the length of the body increasing to accommodate the profiles, resulting in a broadly constant drag for a given width. In the case of the Blunted Wedge profile there even appears to be a small downward trend as the volume increases. It should be noted that the drag values do start to rise, but this appears to be as a result of approaching the boundary of the data set. If the data set were larger, the plateaued trend would likely continue.

The Blunted Wedge is once again the most effective at reducing the drag. As seen in Table 4.24 this is initially achieved using the sharpest profile, with the length of the body doubling compared to the equally constrained cuboid. Above 0.35m^3, however, the morphology changes with the tail profile becoming a prominent feature. The key reason for this is that the optimizer has reached the maximum length of the satellite body in the data set, that being 1.5m. So for the volume to continue to grow while minimizing the drag, but without further increasing the length, secondary features such as the tail profile have to be considered. Similar observations can be made for the Double-Conic and Rounded-Conic profiles, where deviation from the sharp profile only occurs once the body has reached the length constraint of the data set itself.

134
What is interesting about these borderline profiles is that they still achieve significant improvement in the drag force. The 0.5m wide Blunted Wedge for instance still attains a 14% improvement on the equivalent cuboid for an internal volume of 0.50m³ and has only a marginal increase in the drag over the 0.45m³. In fact, as can be seen from Figure 4.35, the rise in the drag for the 0.5m fixed-width profile only really occurs for volumes greater than 0.60m³, at which point the optimizer is hard up against the edge of the data set on multiple parameters. Similar observations can be made of the other profiles and width constraints. This was observed in the previous section as well, a family of near-perfect profiles that achieved the majority of the drag reduction with a morphology that is arguably more usable than the sharp conic profiles.

Unlike the other three profiles, the Elliptical profile did not grow to twice the length of the cuboid to achieve its minimum drag, but it also did not perform as well. The bulbous nature of the profile is effectively like a cone with a non zero nose radius, so has more internal volume for the same length nose. It was hoped that this would result in more usable internal volume, however, it also means a significant proportion of this frontal area is not contributing to reducing the drag. This explains why the optimizer selects for profiles that include a tail long before the edge of the data set is reached. Like the borderline Blunted Wedges, these Elliptical profiles arguably have a larger proportion of the internal volume that is usable compared to the sharp conic profiles. However, seeing as a 0.5m² Blunted wedge achieves a ‘similar’ morphology using flat panels while also
achieving a higher reduction in drag, it is unlikely that an Elliptical profile would be the best choice. This is further compounded when considering that the Elliptical profiles would be more difficult to construct compared to Blunted Wedge due to the curved panels.

4.5.4 Summary

In this section, candidate satellite geometries were presented for a prescribed satellite volume. To help explore the design, space the length and width of the body were constrained separately to simulate the dimensional needs of the internal subsystem. It was shown that when the length of the body was constrained, only a marginal decrease in the drag could be achieved using any of the profiles since the body has to increase its frontal area to accommodate the profile. By contrast, when the width is constrained, there was little variation in the drag experienced by each of the profiles. This was possible because the profiles were able to grow in length, demonstrating that even simply attaching a profile to the front of the profile would reduce the drag overall.

Of the profiles, the Blunted Wedge performed the best overall. In the length-constrained scenario, the profile grew by 5-10% and adopted a wide nose and tail which achieved a reduction in the drag of between 2% and 6% compared to the comparable cuboid. In the width-constrained scenario, the profile adopted the sharp conic form, doubling in length and achieved as much as 21% reduction for the narrowest profile.

4.6 Future Work

To carry out the work for this chapter several assumptions had to be made about the design space and in particular the geometries of the satellite. First and foremost, the profiles examined here were all two dimensional. As a first step, it would be useful to extend the analysis to three-dimensional geometries to see if the relationships identified in this research still hold. This would also allow the examination of the choice of cross-section and its impact on the nose and tail profiles.

Finally, the work presented in this chapter focused exclusively on the drag of the central body. Other subsystems such as the solar panels have not been considered as part of the study. Given their large surface area, solar panels represent a significant contribution to the drag of the satellite and represent a vital link in the drag-power optimization loop. Therefore, further work assessing the interaction between the central body and the solar panels is warranted. This could then lead to a broader systems analysis with the geometry of the satellite as part of the loop.
4.7 Conclusion

This chapter has looked at work assessing the effectiveness of altering the geometry of the satellite to reduce its drag, to either prolong its operational life or increase its payload capacity. In data obtained using the DSMC surrogate models, it has been shown that a maximum drag reduction of between 21% to 35% was achievable for the 1:2 and 1:6 profiles respectively, when optimizing purely for minimum drag assuming $\alpha = 0.95$. To achieve this, all profiles (except for the Elliptical profile) opted for a sharp nose geometry, however, this often resulted in a significant loss of internal volume due to the constraints on the geometry, namely a fixed width and length.

By trading off lowering the drag against the internal volume of the profiles using multi-objective optimization a set of Pareto-optimal satellite geometries could be identified. Under these circumstances, the Elliptical profile generally performed the best of the 4 profiles, particularly for higher aspect ratio profiles. This was especially the case when considering the trade-offs, i.e. the maximum reduction in volume for which the reduction in drag is greater than or equal to the volume reduction. In this case, it was seen that for the 1:6 profiles a drag reduction of 26% was achievable with a reduction in the volume of just 22%. This compares to the 1:6 Blunted Wedge profile that achieved a reduction in both the Drag and volume of 22%. It was also generally observed that the ‘nose-only’ profiles did not perform as well as the Blunted Wedge or Elliptical profiles which both featured a tail section. This was the case even when the optimizer was solely optimizing for drag, and the key difference was that the tail section effectively shortened the profiles by turning the rear surfaces away from the flow.

Finally, satellite geometries were presented for when the volume of the satellite is prescribed. To help explore the design space, the length and width of the body were constrained separately, in an attempt to simulate the dimensional needs of internal subsystems. It was shown that when the length of the body was constrained, only a marginal decrease in the drag could be achieved using any of the profiles since the body has to increase its frontal area to accommodate the profile. By contrast, when the width is constrained, there was little variation in the drag experienced by each of the profiles. This was possible because the profiles were able to grow in length, demonstrating that even simply attaching a profile to the front of the profile would reduce the drag overall.

Of all the profiles, the Blunted Wedge performed the best overall. In the length constrained scenario, the profile grew by 5% to 10% and adopted a wide nose and tail which achieved a reduction in the drag of between 2% and 6% compared to the comparable cuboid. In the width constrained scenario, the profile adopted the sharp conic form, doubling in length and achieved as much as 21% reduction for the narrowest profile.
5.1 Introduction

A key observation in Chapter 4 was the link between the gas refill rate behind the satellite and the ideal geometry of the tail. In particular, beyond a certain internal tail angle, the tapered surface was effectively shielded by the body of the satellite and thus had a limited contribution to the drag experienced. This ultimately proved beneficial while optimizing the satellite geometry, especially for those geometries with high aspect ratios. Beyond this, the refill rate also has implications for other aspects of the geometry, in particular the ingress and propagation of ATOX to the internal environment of the satellite.

As outlined in Section 1.2.5, most research into the effects of ATOX focuses on its impact on the external portions of the satellite. For normal LEO this is sufficient, however, at the lower altitudes in VLEO the density of ATOX is much higher (as seen in Section 2.3.3) and the ingress of gases to the satellite’s interior through openings could leave payloads (for example) exposed to the corrosive nature of ATOX which could limit their operational lives.

This chapter sets out to answer three questions: Firstly, how much ATOX can be expected to permeate into the interior of a satellite through an opening on the side and whether this is dependent in any way on the size of the hole and what is the ambient flux level. Secondly, as a result of this infiltration whether there are any hotspots specifically of ATOX and if so what is the expected flux. Finally, whether these values can be predicted either through analytically analysis or empirically data.
5.2 General Method

5.2.1 Workflow

The work presented in this chapter focuses on the influx of ATOX and other gases, into the interior of the satellite. ATOX is incredibly corrosive, so understanding how it propagates into the body of the satellite is essential for identifying ways of protecting delicate instruments. An overview of the work carried out as part of this chapter is provided in Figure 5.1. In general, the work followed two approaches, a simulated approach (left side of Figure 5.1) and an analytical approach (right side of Figure 5.1).

![Figure 5.1: Overview of workflow in Chapter 5](image-url)
In the analytical approach, gaskinetic theory is applied to determine what the expected distribution of the flux should be within the pit. This begins with deriving the angular deviation for gases in a rarefied flow and understanding the properties of their distributions to make predictions about how the flow would behave within the cavity, as will be outlined in Section 5.3. This was followed by performing a second set of derivations for the expected particle flux distributions on the surfaces within the pit in Section 5.3.3.

The second approach involved modelling the rarefied gas dynamics using DSMC simulations. To achieve this a simple rectangular recessed pit was defined, as shown in Section 5.2.2, and the gas dynamics were simulated using the DSMC software SPARTA. As the surface properties of the material are not known, the particle-surface interactions were assumed to be diffuse with a range of thermal energy coefficients ($\alpha_T$) from 0 (specular) to 1 (fully diffuse). The simulations will assume that the satellite body is orbiting in a circular orbit of 200km. The atmospheric conditions at this altitude will be assumed to be constant around the orbit, as outlined in Section 5.2.3.

In Section 5.3.3.4 the analytical results were compared with those derived from the DSMC simulations with fully diffuse particle-surface interactions ($\alpha_T = 1$). This provided feedback on how well the analytical results could predict the fluency on the surfaces as well as identify any anomalies. This will be followed by performing an analysis across the range of $\alpha_T$ in Section 5.4.1 to identify hotspots and determine the depth of penetrations ATOX particles.

5.2.2 The Pit Geometry

To explore the possible effects on the internal environment a simple simulation set-up was devised. This comprised a single rectangular pit, recessed into a wall that ran parallel to the free-stream velocity ($v_\infty$) as can be seen in Figure 5.2. Figure 5.2 also includes the locations of the Forward-facing panel (FFP), Rearward-facing panel (RFP), Pit floor panel (PFP), Forward-facing rim (FFR) and Rearward-facing rim (RFR).
facing panel (FFP), Rearward-facing panel (RFP), Pit floor panel (PFP), Forward-facing rim (FFR) and Rearward-facing rim (RFR). Both the pit’s depth \( (d_p) \) and length \( (l_p) \) were varied between 0.1m and 0.5m. At the lower limits, this is representative of small payload openings and loose connections between adjacent panels. The upper limit is more extreme and may not be strictly representative of what could be feasibly or desirably included on a satellite. However, for Earth observation purposes some payloads may require large apertures to the exterior. These upper limits were therefore chosen to be an inclusive upper boundary rather than reflect the largest possible pit size for a satellite. To help capturing the distribution of particles striking the surfaces the walls of the pit were each divided into 20 equal segments.

To help facilitate the simulation, the front edge of the pit is set back from the forward inflow boundary. This lets the flow become established and interact with itself, (thus allowing some particle-particle interactions to take place) before interacting with the geometry of the pit. The surface upstream of the pit was extended to the forward inflow boundary to reduce the effect of particle pile up around the upstream section of the simulation.

As described above, the pit is intended to represent a diverse set of situations from panel joints to payload openings. Larger openings may feature equipment mounted to the various internal surfaces, however, no attempt will be made to model these. The goal here is to achieve a generalized understanding of the gas ingress. Thus a simple rectangular pit was chosen.

### 5.2.3 Atmospheric Conditions

For this analysis, it will be assumed that the atmospheric conditions remain constant for the chosen altitude using a global average of the atmospheric properties taken during high solar activity on the 1st March 2000 (see Section 2.3.4). The gas species included in the model were: Oxygen \( (O_2) \), Nitrogen \( (N_2) \), Atomic Oxygen \( (O_1) \), Atomic Nitrogen \( (N_1) \), Argon \( (Ar) \), Helium \( (He) \) and Hydrogen \( (H) \). A summary of the derived properties used is presented in Table 5.1 for a 200km circular orbit.

Table 5.1: Properties of the atmosphere at 200km on the 1st March 2000 at 00:00:00 UTC as given by the spherically weighted mean of NRLMSISE-00 model\(^{[30]}\)

<table>
<thead>
<tr>
<th>Altitude [km]</th>
<th>Number Density (m^{-3})</th>
<th>Temperature [K]</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>(1.06 \times 10^{16})</td>
<td>1063</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Composition [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atomic Oxygen ( (O_1) )</td>
</tr>
<tr>
<td>Nitrogen ( (N_2) )</td>
</tr>
<tr>
<td>Oxygen ( (O_2) )</td>
</tr>
<tr>
<td>Helium ( (He) )</td>
</tr>
<tr>
<td>Atomic Nitrogen ( (N_1) )</td>
</tr>
<tr>
<td>Argon ( (Ar) )</td>
</tr>
<tr>
<td>Hydrogen ( (H) )</td>
</tr>
</tbody>
</table>
5.3 Directional probability distribution of Gasses

5.3.1 Introduction

This section will deal with the derivation of the analytical solution for the flux distribution in a simple rectangular pit. The first part of this section, Section 5.3.2, will outline the derivation of the angular probability distribution for gas species in a rarefied flow. Using this knowledge, generalized equations for the particle flux on the surface of panels within a pit mounted on the side of the satellite will be derived in Section 5.3.3. This will subsequently be compared to DSMC simulations of a 0.5m wide rectangular pit in a 200km orbit with diffuse particle-surface interactions.

5.3.2 The Angular Distribution of Gas Species

5.3.2.1 Velocities of a Particle

The velocity of particles in a rarefied gas ($v$) has two key components, the mean velocity of the bulk fluid ($v_\infty$) and the internal velocity ($\tilde{v}$) of the fluid. If the internal velocity $\tilde{v}$ follows a known distribution (usually determined by its temperature), then for a homogeneous fluid the particles passing through a given point in space will follow an angular probability distribution ($F(\theta)$), as outlined in Figure 5.3. As the flow in VLEO is non-continuous (see Section 2.4), it is the random lateral motion of the particles that brings them in contact with the sides of the satellite and through any opening along them. This section aims to obtain the distribution $F(\theta)$ for gases in a rarefied flow and to examine the angular distribution of species present in VLEO.

![Figure 5.3: Diagram showing the velocity vectors and the projected angular probability distribution function $F(\theta)$](image)

143
5.3.2.2 Distribution of Velocities

The direction of the bulk gas velocity \( \mathbf{v}_\infty \) is fixed by the macroscopic circumstances of the flow, in this instance a satellite passing through a region of VLEO. The magnitude and direction of the internal velocities \( \mathbf{\tilde{v}} \) can be characterized by a thermal distribution function. Assuming that within the moving frame of the bulk fluid the particles have achieved kinematic equilibrium, then the Maxwell-Boltzmann distribution can be used to determine these velocities\[45\].

In 3 dimensions the Maxwell-Boltzmann distribution is given by Equation 5.1. \( u, v \) and \( w \) in Equation 5.1 are elements of the velocity vector \( \mathbf{v} \) and represent velocities in the \( x, y \) and \( z \) axis respectively. Here the Maxwell-Boltzmann distribution is expressed in terms of the most probable thermal speed \( c' \) which is given by Equation 5.2, where \( m \) is the mass of the particle of the species, \( T \) is the temperature of the gas and \( k \) is the Boltzmann constant.

Equation 5.4 therefore becomes

\[
f(\mathbf{v})dudvdw = \left( \frac{1}{\pi c'^3} \right)^\frac{3}{2} \exp\left( -\frac{|\mathbf{v}|^2}{c'^2} \right)dudvdw
\]

\[
c' = \sqrt{\frac{2kT}{m}}
\]

\[
s = \frac{v}{c'} \quad s_\infty = \frac{v_\infty}{c'} \quad s_r = \frac{v_r}{c'} \quad s_w = \frac{w}{c'}
\]

At this stage, it is also useful to introduce the molecular speed ratio \( s \), given by Equation 5.3. Thus for total velocity, the Maxwell-Boltzmann Equation 5.1 can be rewritten

\[
f(\mathbf{v})dudvdw = c'^3f(s)dsdsdsw = \left( \frac{1}{\pi} \right)^\frac{3}{2} \exp\left( -|s - s_\infty|^2 \right)dsudsvds_w
\]

5.3.2.3 Polar Form

The next step is to convert the distribution in Equation 5.4 to a polar form with relation to the direction of \( \mathbf{v}_\infty \) to derive the angular distribution. Typically blockages in the flow will end on a straight edge with a finite length, such as the rear edges of the satellite or the rim of a recessed pit. Therefore it was decided to express the bulk velocity vector \( \mathbf{v} = [u, v, w] \) in terms of the cylindrical coordinates \([v_r, \theta, w]\) where \( v_r \) is the 2D velocity vector \([u, v]\) and \( w \) is parallel to the rim of the pit and normal to the plane of the figure (see Figure 5.2), \( \theta \) is the direction of \( v_r \) with respect to \( v_\infty \).

\[
u = |v_r|\cos(\theta) \quad u = |v_r|\sin(\theta) \quad w = w
\]

\[
s_u = |s_r|\cos(\theta) \quad s_v = |s_r|\sin(\theta) \quad s_w = s_w
\]

Equation 5.4 therefore becomes

\[
f(\mathbf{v})dudvdw = \left( \frac{1}{\pi} \right)^\frac{3}{2} \exp\left( 2|s_r||s_\infty|\cos(\theta) - |s_r|^2 - |s_\infty|^2 - s_w^2 \right)|s_r|ds_r|d\theta|ds_w
\]
5.3. DIRECTIONAL PROBABILITY DISTRIBUTION OF GASSES

Figure 5.4: Angular probability distribution of the molecular flux \( F(\theta) \) relative to the bulk fluid velocity \( \vec{u}_\infty \)

where \( s_\infty \) is the molecular speed ratio of the bulk velocity \( \vec{u}_\infty \) (Equation 5.3). The distribution with respect to the angle \( \theta \), \( F(\theta) \), can then be obtained by integrating Equation 5.7 with respect to \( ds \) and \( d\theta \), therefore:

\[
F(\theta) = \frac{1}{2\pi} \left( \sqrt{|s_\infty| \cos(\theta) \exp(-|s_\infty|^2 \sin^2(\theta)(\text{erf}(|s_\infty| \cos(\theta)) + 1) + \exp(-|s_\infty|^2))} \right)
\]  

Equation 5.8 represents the angular probability distribution of the molecular flux \( F(\theta) \) relative to the bulk fluid velocity \( \vec{u}_\infty \). The distribution \( F(\theta) \) is shown in Figure 5.4 for a selection of \(|s_\infty|\). More details on the derivation of Equation 5.8 can be seen in Appendix C.1

5.3.2.4 Refill Rate

In the non-viscous environment of the upper atmosphere, the spread of the probability distributions of molecular flux in Figure 5.4 determines how quickly the fluid will refill. As can be seen from this figure, the distribution appears to approach the form of a normal distribution. Indeed as is demonstrated in Appendix C.1.6 by applying the small-angle approximation an estimate of the angular standard deviation \( \sigma_{AD} \) can be made in the form of Equation 5.9.

Figure 5.5 provides a comparison between \( \sigma_{AD} \) and the numerically calculated standard deviation of \( F(\theta) \) for a range of \(|s_\infty|\). As can be seen from this figure, \( F(\theta) \) converges towards a normal distribution as \(|s_\infty|\) increases.

The refill rate of rarefied gases was first described by Koppenwaller et al. in their work on concave corners, in which they introduced the concept of the molecular Mach angle \( \mu \) given by Equation 5.10. Conceptually, \( \mu \) is the angle a particle would take if its internal velocity \( \vec{v}_i \) was equal to...
CHAPTER 5. INTERNAL ENVIRONMENT IN VLEO

Figure 5.5: Comparison of $\sigma_{AD}$ and $\mu$ with the numerical calculated standard deviation of $F(\theta)$ the most probable speed $c'$ in a direction perpendicular to $v_\infty$. As with $\sigma_{AD}$, the $\mu$ represents the extent to which the particle flux deviates from $v_\infty$. A comparison of $\mu$ and $\sigma_{AD}$ are shown in Figure 5.5 for a range of $|s_\infty|$, along with a result for the standard deviation calculated numerically from the angular distributions in Figure 5.4.

$$\sigma_{AD} = \sin^{-1} \left( \frac{1}{|s_\infty| \sqrt{2}} \right) \tag{5.9}$$

$$\mu = \tan^{-1} \left( \frac{1}{|s_\infty|} \right) \tag{5.10}$$

The key result that Koppenwaller et al. demonstrated was that in a concave corner, some degree of shading of one or both corner faces would occur if the local angle of incidence to either face was less than 2.5$\mu$ for that gas species. This is demonstrated in Figure 5.6, which shows a schematic of the three possible shading scenarios for a concave corner. [92]. Above 2.5$\mu$, the effects of shading were comparatively negligible and so the faces could be assumed to be fully immersed in the flow, as long as there were no other obstructions.

Figure 5.6: Shielding effect criteria for concave corners from Koppenwaller et al. [92]
5.3. DIRECTIONAL PROBABILITY DISTRIBUTION OF GASSES

Figure 5.7: Proportion of particles contained within a cone with a half angle of $3\sigma_{AD}$ and $2.5\mu$. $3SD$ cone from the numerical standard deviation included for comparison.

From this, it might be inferred that the angle given by $2.5\mu$ can be used to provide an estimate of the maximum rate at which a gas will refill a region following a blockage. Applying this to the angular distribution, it can be seen in Figure 5.7 that a cone with a half-angle of $2.5\mu$ would contain 99.9% of all particles. Similarly, a cone with a half-angle of $3\sigma_{AD}$ would contain roughly 99.7% of all particles as would be expected from the $3\sigma$ rule.

While a useful metric, $\mu$ has no specific link to the angular distribution of the particles. The most probable speed, for instance, forms a ring of velocities around $v_\infty$ (see Figure 5.3), so it would be incorrect to describe $\mu$ as the most probable angle. As can be seen from Figure 5.3, $\sigma_{AD}$ provides a good estimate of the standard deviation of the angular distribution, especially for high values of $\vert s_\infty \vert$. Since the angular distribution (Equation 5.8) does approximate a normal distribution at high $\vert s_\infty \vert$ (see Appendix C.1.6), $\sigma_{AD}$ directly describes the nature of the flux distribution. The key driver of this analysis is knowing the inbound particle density and for that reason, the angular deviation ($\sigma_{AD}$) is more useful and will therefore be the preferred metric in later sections.

5.3.2.5 The Angular Distribution of Gas Species

In the previous sections, an equation for the angular distribution of the molecular flux in a rarefied gas was derived. In this section, the distribution is applied to the primary gas species present in VLEO that being Atomic Oxygen ($O_1$), Nitrogen ($N_2$), Molecular Oxygen ($O_2$), Helium ($He$), Atomic Nitrogen ($N_1$), Argon ($Ar$) and Hydrogen ($H$). Figure 5.8 shows the 2D angular probability distribution of the gas species at an altitude of 200km as given by Equation 5.8. $v_\infty$ is assumed to be the velocity of the satellite in a circular orbit at that altitude. The atmospheric conditions used to compute these distributions at this altitude are provided in Table 5.1. By multiplying each angular distribution by the species’ number density, an estimate of the angular fluencies can be made as shown in Figure 5.9.
Table 5.2: Mean free speed ratio for the gas species to be simulated in subsequent DSMC simulations and their associated angular deviations $\sigma_{AD}$ and molecular Mach angles $\mu$

<table>
<thead>
<tr>
<th>Gas Species</th>
<th>Composition [%]</th>
<th>$s_\infty$ [-]</th>
<th>$\sigma_{AD}$ [°]</th>
<th>$\mu$ [°]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atomic Oxygen ($O_1$)</td>
<td>56.20</td>
<td>7.41</td>
<td>5.48</td>
<td>7.68</td>
</tr>
<tr>
<td>Nitrogen ($N_2$)</td>
<td>39.89</td>
<td>9.80</td>
<td>4.14</td>
<td>5.82</td>
</tr>
<tr>
<td>Oxygen ($O_2$)</td>
<td>2.97</td>
<td>10.48</td>
<td>3.87</td>
<td>5.45</td>
</tr>
<tr>
<td>Helium ($He$)</td>
<td>0.59</td>
<td>3.71</td>
<td>11.00</td>
<td>15.10</td>
</tr>
<tr>
<td>Atomic Nitrogen ($N_1$)</td>
<td>0.28</td>
<td>6.93</td>
<td>5.85</td>
<td>8.21</td>
</tr>
<tr>
<td>Argon ($Ar$)</td>
<td>0.06</td>
<td>11.71</td>
<td>3.46</td>
<td>4.88</td>
</tr>
<tr>
<td>Hydrogen ($H$)</td>
<td>0.01</td>
<td>1.86</td>
<td>22.35</td>
<td>28.27</td>
</tr>
<tr>
<td>Mixture</td>
<td>100.00</td>
<td>8.53</td>
<td>4.76</td>
<td>6.69</td>
</tr>
</tbody>
</table>

Figure 5.8: Angular probability distribution of the species in the atmosphere at an altitude of 200km under the condition provided in Table 5.1

$H$ and $He$, as the lightest elements, have the greatest $\sigma_{AD}$ of around 22.35° and 11.00° respectively. This implies that they would be the fastest to refill the void behind a satellite or enter a recessed pit. $H$ and $He$ also have the lowest speed ratios of 1.86 and 3.71 respectively. As was outlined in the section above, $\sigma_{AD}$ as calculated using Equation 5.9 is only valid when $|s_\infty| > 5$ which does raise questions about the accuracy of these $\sigma_{AD}$. However, they also make up a small fraction of the overall gas mixture and therefore do not have a large impact on the mixtures’ angular distribution as a whole or its $\sigma_{AD}$. As can be seen from Table 5.2 and Figure 5.9, the bulk fluid is dominated by $O_1$ and $N_2$ which comprise 56.20% and 39.89% of the gas mixture respectively. The two species have a similar $\sigma_{AD}$ equal to 5.48° and 4.14° respectively so, given their dominance, the $\sigma_{AD}$ of the gas mixture’s angular distribution sits between them at 5.02°.
5.3. DIRECTIONAL PROBABILITY DISTRIBUTION OF GASSES

Figure 5.9: Angular number density distribution of the various species in the atmosphere at an altitude of 200km under the condition provided in Table 5.1

5.3.3 Flux Distribution on a Wall

5.3.3.1 Introduction

In the previous section, the angular probability distribution for gas species in a rarefied flow was derived in Equation 5.8. This was useful for understanding how the particles diverge from $\mathbf{v}_\infty$ in a rarefied gas. This section will extend this understanding by outlining analytical solutions for the particle flux on the surface of panels within a pit mounted on the side of the satellite. These solutions will then be compared to a set of DSMC results to verify the equations.

5.3.3.2 Flux on a Wall

For a simple flat plate presented to a rarefied flow under steady-state conditions, the particle flux $\Omega_{N_s}^X$ for a given species $S$ at a point $X$ along the surface is given by Equation 5.11, where $n_s$ is the number density of the species in the free-stream, $\mathbf{v}_n$ is the velocity normal to the surface and $f(\mathbf{v})$ is the Maxwell-Boltzmann distribution for the velocity vector $\mathbf{v}$ of the species.

$$\Omega_{N_s}^X = n_s \iiint_{\text{inbound}} f(\mathbf{v})|\mathbf{v}_n| \, du \, dv \, dw \quad (5.11)$$

If the surface is fully exposed to the free-stream (as seen in Figure 5.10a), Equation 5.11 can be integrated over all inbound velocities to determine the particle flux at the surface. However, in the case of a simple recessed pit, the particles in the bulk fluid can only reach the surfaces within the pit through the opening at the rim (as seen in Figure 5.10b). This means that for a particle to interact with a surface element $X$ inside the pit, the particle must have an incoming radial velocity.
\( \mathbf{v}_r \) with an incident angle between \( \eta_f \) and \( \eta_r \), the angles from X to the FFR and RFR of the pit respectively (see Figure 5.11).

As in the previous section, it is useful to express the velocity \( \mathbf{v} = [u, v, w] \) in terms of cylindrical coordinates \([r, \theta, w]\) (Equation 5.6). As before, \( \mathbf{v}_r \) is the 2D radial velocity vector \([u, v]\) and \( w \) is parallel to the rim of the pit and normal to the plane of Figure 5.10a. In this case, \( \theta \) is the direction of vector \( \mathbf{v}_r \) with respect to the normal of surface \( X \). Given this, the velocity normal to the surface \( \mathbf{v}_n = \mathbf{v}_r \cos(\theta) \). Equation 5.11 therefore becomes

\[
\Omega^N_X = n_s c' \int_{\eta_f}^{\eta_r} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\mathbf{v}) |\mathbf{v}_r| |\mathbf{v}_n| d\mathbf{v}_r d \mathbf{v}_n d\theta
\]  

Equation 5.12

\( f(\mathbf{v}) \) was previously defined in Section 5.3.2 with respect to the cylindrical coordinates \([r, \theta, w]\) by Equation 5.7, where \( \theta \) was the direction of \( \mathbf{v}_r \) with respect to the bulk velocity \( \mathbf{v}_\infty \). Inserting this into the Equation 5.12 and expressing the velocities in terms of the molecular speed ratio \( s \) yields

\[
\Omega^N_X = n_s c' \int_{\eta_f}^{\eta_r} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left( \frac{1}{\pi} \right)^\frac{3}{2} \exp \left( \frac{1}{2} |\mathbf{s}_\infty|^2 - \frac{1}{2} |\mathbf{s}_\infty|^2 - s_w^2 \right) |\mathbf{s}_n| |\mathbf{s}_r| d\mathbf{v}_r d \mathbf{v}_n d\theta
\]

Finally, integrating this over \( d\mathbf{v}_n \) and \( d|\mathbf{s}_r| \) produces

\[
\Omega^N_X = n_s \int_{\eta_f}^{\eta_r} \left( \frac{c'}{2\pi} \right) \left( \frac{\sqrt{\pi}}{2} \right) \left( \frac{1}{2} \right) \exp \left( -|\mathbf{s}_\infty|^2 \sin^2(\theta) \right) \left( \text{erf}(\sqrt{|\mathbf{s}_\infty|^2 \cos(\theta)}) + 1 \right) |\mathbf{s}_\infty| \cos(\theta) \exp(-|\mathbf{s}_\infty|^2) \cos(\theta) d\theta
\]

Equation 5.14 is the general form for the flux on a surface in the pit as a result of the gases entering the pit from the exterior. This equation does not account for reflections from other surfaces, though an approximation can be made by applying this equation recursively with a reflection model.

150
5.3. DIRECTIONAL PROBABILITY DISTRIBUTION OF GASSES

5.3.3.3 Forward-facing panel (FFP)

To perform a comparison with DSMC simulations it is necessary to rewrite Equation 5.14 for the surface in a simple recessed pit and in particular the FFP. It can be seen from Figure 5.11 that for a point \( X \) on the surface, the incoming velocity vectors must lie between \( \theta_X \) and 90° (\( \pi/2 \)). Equation 5.14 therefore becomes

\[
\Omega_{\mathcal{N}}^{X,s} = n_s \int_{\theta_X}^{\pi/2} \left( \sqrt{\pi} \left( |s_\infty|^2 \cos^2(\theta) + \frac{1}{2} \right) \exp(-|s_\infty|^2 \sin^2(\theta)) (\text{erf}(|s_\infty| \cos(\theta)) + 1) 
+ |s_\infty| \cos(\theta) \exp(-|s_\infty|^2) \cos(\theta) \right) d\theta \tag{5.15}
\]

This can be integrated over \( d\theta \) to

\[
\Omega_{\mathcal{N}}^{X,s} = \frac{n_s c'}{4\pi} \left( (\text{erf}(|s_\infty|) + 1)(\pi|s_\infty| - \text{erf}(|s_\infty| \sin(\theta_X))) 
- \sqrt{\pi} \sin(\theta_X) \exp(-|s_\infty|^2 \sin^2(\theta_X)) + \sqrt{\pi} \sin(\theta_X) \exp(-|s_\infty|^2) \right) 
\]

\[
\frac{-|s_\infty| \exp(-|s_\infty|^2)}{2}(2\theta_X + \sin(2\theta_X) - \pi) \tag{5.16}
\]

When \( |s_\infty| > 5 \) Equation 5.16 can be simplified and rearranged to form Equation 5.17 (more detail on this simplification is given in Appendix C.2). As can be seen from this new equation, the flux on the FFP can be expressed in terms of the maximum expected flux on surfaces perpendicular to the flow \( N_\infty \) (Equation 5.18) and can be obtained by integrating over all incoming velocity. Furthermore, Equation 5.17 also approximates the cumulative distribution function of a normal distribution. This is broadly to be expected since it was shown in Section 5.3.2 that the angular distribution of particles can be approximated to a normal distribution with respect to \( \sin(\theta) \). A cumulative distribution function is simply the probability density function integrated between \(-\infty\) and \( x \) which is similar to how Equation 5.17 was obtained.

\[
\Omega_{\mathcal{N}}^{X,s} = N_\infty \left( \frac{1}{2} (1 - \text{erf}(|s_\infty| \sin(\theta_X))) - \frac{\sin(\theta_X) \exp(-|s_\infty|^2 \sin^2(\theta_X))}{2|s_\infty| \sqrt{\pi}} \right) \tag{5.17}
\]

\[
N_\infty = n_\infty |v_\infty| = n_\infty c' |s_\infty| \tag{5.18}
\]

Figure 5.11: Diagram showing the inflow to a point \( X \) on the surface of the FFP in a rectangular recessed pit
5.3.3.4 The Wall Distribution of Gas Species in a simple Pit

The next step is to compare the analytical equation to the results of the DSMC simulations described in Section 5.2 for the simple recessed pit. The graph in Figure 5.13 shows the flux distribution for each species (as well as the gas mixture as a whole) along the FFP of a 0.5m wide recessed pit on the side of a satellite (as described in Section 5.2.2) in a 200km circular orbit. In each case, the predicted wall distribution using Equation 5.17 and the results of a set of DSMC simulations are shown. For the DSMC simulations, the surfaces of the satellite were assumed to interact with the particles diffusely with $\alpha = 1$.

Figure 5.13 also shows the maximum expected flux ($N_\infty$) and the ambient flux ($\tilde{N}_\infty$) for each species. The $N_\infty$ assumes the panel is fully exposed to the incoming fluid and can be approximated using Equation 5.18, while $\tilde{N}_\infty$ is taken to be the average flux on the forward and aft outer walls that run parallel to the flow in the DSMC simulation. The figures also show the projected first, second and third $\sigma_{AD}$ cones for the angular distribution as well as the 2.5$\mu$ cone (discussed in Section 5.3.2) projected from the start of the opening down into the pit.

From the graphs in Figure 5.13 it can be seen that, in general, the flux distribution from the DSMC simulations matches the predicted trend of Equation 5.17. This is partially to be expected since DSMC methods typically use the Maxwell-Boltzmann distribution to set the initial velocity of the simulated particles. This is certainly the case for SPARTA [56], the DSMC tool that was used to perform the simulations in this case. Undeniably a better contrast would be to compare these results to a set of lab-based tests, however, the facility that could conduct this experiment is still under development. This could therefore be a good avenue for future work.

Despite the similarity in the trend, it was observed that for each species ambient flux in the pit ($\tilde{N}_{\text{pit}}$) is non-zero, which is easily visible at the lowest regions of the FFP (rightmost end of the graphs in Figure 5.13). $\tilde{N}_{\text{pit}}$ is predominantly the result of slow-moving particles captured by the pit following diffuse particle-surface interactions with the pit walls. Identifying this is critical to determining to what extent the flux on the surfaces of a confined space within the satellite can be predicted by analytical methods without resorting to more complex tools such as DSMC and lab testing and this will be discussed in more detail below.

As would be expected for a recessed pit in a rarefied flow, the maximum flux on the FFP was observed at the rim of the pit. This was observed for all gas species as can be seen in Figure 5.13. From Figure 5.13 and Table 5.3 it can be seen that this peak flux is approximately half the $N_\infty$ for the given species. The physical explanation for this is that only particles whose velocity ($v$) bring them into the pit will have an interaction with the interior pit surfaces. As $v_\infty$ is parallel to the outer surface walls and based on the distribution in Figure 5.8, it can be expected that approximately half the particles will have an orthogonal velocity component that meets these criteria. The other half of the distribution would be blocked by the structure of the satellite.

A similar explanation can be derived from Equation 5.3.2 where it is shown that the limits on the integral for a point near the rim of the FFP would be 0 and $\pi/2$ radian. This is a quarter of the interval of the angular distribution but accounts for approximately half the particles in the stream,
as the distribution is concentrated to a narrow distribution around 0 degrees when $|s_\infty|$ is large. It can therefore be expected that the flux on the rim of a recessed pit will be half that seen on the front of the satellite ($N_\infty$), assuming the pit is perpendicular to the flow.

A similar relationship is observed for the ambient flux $N_{pit}$, in that it is approximately half $N_\infty$ as can be seen in Table 5.3. This second result was unexpected as there is no immediate mechanism that links these two fluency values. For instance, $N_\infty$ is simply the result of the orthogonal velocity bringing the gas particles in contact with these surfaces instead of down into the pit itself. The ambient flux at the bottom of the pit on the other hand is the result of the particles captured following diffuse interactions with the FFP. It might be expected that over time, the ambient level may increase as more particles are captured and thus accumulate, however, this is not observed in the longer duration DSMC simulations. Figure 5.12 shows the ratio of $N_{pit}$ to $N_\infty$ from the DSMC simulations over time for the gas species. While there is some fluctuation, $N_{pit}$ quickly settles and then remains approximately $N_\infty$ with no observed growth over time. Furthermore, the average number density across the pit appears to stabilize around the free stream number density indicating that an equilibrium has been reached.

In addition to the maximum expected flux at the rim of the pit, it is also important to know the extent of its associated hotspot and in the case of a hotspot near the rim of a recessed pit, the extent can be defined in terms of the refill rate of the gas. As will be discussed in subsequent sections, this is vital for identifying areas that will be most affected by ATOX, not just near the rim but also following multiple reflections in instances where the particle surface interactions are more specular. As can be seen from Figure 5.13 the absolute extent of the hotspot can either be described by the $3\sigma_{AD}$ or 2.5$\mu$ refill cone. In this instance, the predicted flux on the wall relative to $N_\infty$ is 0.16% and 0.03% respectively. The advantage of $\sigma_{AD}$ here, as was also discussed in Section 5.3.2 is that the flux distribution across the FFP closely resembles that of a cumulative frequency distribution for a normal distribution. This means that $\sigma_{AD}$ can be used to describe how the particles will be distributed along that surface.

Table 5.3: Comparison of maximum and ambient fluencies for each gas species on the FFP of a simple rectangular pit at 200km altitude Table 5.1

<table>
<thead>
<tr>
<th>Gas Species</th>
<th>Max Flux $N_\infty$ $[m^{-1}]$</th>
<th>Analytical $N_\infty$ $[m^{-1}]$</th>
<th>DSMC $N_\infty$ $[m^{-1}]$</th>
<th>Ambient Flux $N_{pit}$ $[m^{-1}]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atomic Oxygen ($O_1$)</td>
<td>$4.73 \times 10^{19}$</td>
<td>$2.37 \times 10^{19}$</td>
<td>$2.32 \times 10^{19}$</td>
<td>$1.85 \times 10^{18}$</td>
</tr>
<tr>
<td>Nitrogen ($N_2$)</td>
<td>$3.36 \times 10^{19}$</td>
<td>$1.68 \times 10^{19}$</td>
<td>$1.62 \times 10^{19}$</td>
<td>$1.01 \times 10^{18}$</td>
</tr>
<tr>
<td>Oxygen ($O_2$)</td>
<td>$2.50 \times 10^{18}$</td>
<td>$1.25 \times 10^{18}$</td>
<td>$1.21 \times 10^{18}$</td>
<td>$6.92 \times 10^{16}$</td>
</tr>
<tr>
<td>Helium ($He$)</td>
<td>$4.93 \times 10^{17}$</td>
<td>$2.46 \times 10^{17}$</td>
<td>$2.53 \times 10^{17}$</td>
<td>$3.77 \times 10^{16}$</td>
</tr>
<tr>
<td>Atomic Nitrogen ($N_1$)</td>
<td>$2.34 \times 10^{17}$</td>
<td>$1.17 \times 10^{17}$</td>
<td>$1.11 \times 10^{17}$</td>
<td>$9.58 \times 10^{15}$</td>
</tr>
<tr>
<td>Argon ($Ar$)</td>
<td>$5.43 \times 10^{16}$</td>
<td>$2.71 \times 10^{16}$</td>
<td>$2.51 \times 10^{16}$</td>
<td>$1.27 \times 10^{15}$</td>
</tr>
<tr>
<td>Hydrogen ($H$)</td>
<td>$1.11 \times 10^{16}$</td>
<td>$5.56 \times 10^{15}$</td>
<td>$7.35 \times 10^{15}$</td>
<td>$1.12 \times 10^{15}$</td>
</tr>
<tr>
<td>Mixture</td>
<td>$8.42 \times 10^{19}$</td>
<td>$4.21 \times 10^{19}$</td>
<td>$4.10 \times 10^{19}$</td>
<td>$2.97 \times 10^{18}$</td>
</tr>
</tbody>
</table>
5.3.4 Summary

In this section, analytical solutions for the particle flux on surfaces within a simple pit were derived. The first part of this section, Section 5.3.2, outlined the derivation of the angular probability distribution for gas species in a rarefied flow, Equation 5.8. It was demonstrated that the angular distribution can approximate a normal distribution when $|s_\infty| > 5$ and from this, the angular distribution $\sigma_{AD}$ was defined. It was shown that $\sigma_{AD}$ could be used to estimate the distribution of particles in a flow and by extension the proportion of particles that fall within a given dispersion cone angle. In the second part of this section, Section 5.3.3, equations for the particle flux on the surfaces within a pit mounted on the side of the satellite were derived. These were a general case in terms of Equation 5.14 and the more specific case of Equation 5.16 for an FFP in a simple rectangular pit. The specific case was compared to DSMC simulations of a 0.5m wide rectangular pit in a 200km orbit with diffuse particle-surface interactions ($\alpha_T = 1$). It was shown that despite a non-zero $\hat{N}_{\text{pit}}$, the trend of the DSMC data matched that of the analytical result and produced similar predictions for the max flux near the rim. It was also shown that under the diffuse particle-surface conditions the maximum flux in the pit and $\hat{N}_{\text{pit}}$ would be approximately half their respective counterparts experienced on the exterior surfaces. It was also shown that $\sigma_{AD}$ could be used to identify the extent of the hotspot near the FFP and could thus be used to provide an estimate of the fluencies on the FFP.
5.3. DIRECTIONAL PROBABILITY DISTRIBUTION OF GASSES

(a) Atomic Oxygen ($O_1$)

(b) Nitrogen ($N_2$)

(c) Oxygen ($O_2$)

(d) Helium ($He$)

(e) Atomic Nitrogen ($N_1$)

(f) Argon ($Ar$)

(g) Hydrogen ($H$)

(h) Mixture

Figure 5.13: Distribution of particle flux on the FFP of a recessed pit at 200km in altitude
5.4 Atomic Oxygen in a simple pit

A major concern that arises from having an opening in the side of the satellite, particularly for sensitive equipment and payloads, is the inflow of ATOX. As was seen in Section 5.3, under diffuse conditions a hotspot of high flux forms near the rim of the pit on the FFP (see Figure 5.2). In this section, the extent and location of this and other hotspots will be explored within the context of the simple rectangular pit for different \( \alpha_T \) and under the atmospheric conditions at an altitude of 200km (see Table 5.1).

5.4.1 Atomic Oxygen Hot spots

5.4.1.1 General Overview

The graphs in Figures 5.14, 5.15 & 5.16 show the distribution of ATOX for simple rectangular pits with a 0.1m, 0.3m and 0.5m wide openings respectively. The distribution of ATOX on the Forward-facing panel (FFP) is presented on the left of these figures while the Rearward-facing panel (RFP) are presented on the right with each row representing an \( \alpha_T \) of 0, 0.5 and 1 respectively.

Included in the figures are \( N_\infty \), which were estimated using Equation 5.18 for ATOX at 200km altitude. \( N_\infty \) for ATOX as well as the maximum flux values along the FFP and RFP are given in Table 5.4. Also included on the graphs in Figures 5.14, 5.15 & 5.16 are the ambient fluencies on the exterior of the satellite (\( N_\infty \)). This is taken to be the mean flux on the surfaces forward and aft of the pit that lies parallel to the direction of the bulk flow (\( v_\infty \)). These values are also shown in Table 5.4. Finally, the graphs display the estimated extent of the incoming dispersion cone using the value of \( \sigma_{AD} \) for ATOX as well as the predicted locations of subsequent hotspots.

As can be seen from Figures 5.14, 5.15 & 5.16 in all cases the maximum flux is experienced at the top of the FFP. This is the first point at which the fluid interacts directly with the surfaces of the pit and, as can be seen from the figures, the flux is about half \( N_\infty \) (the horizontal dashed line in Figures 5.14, 5.15 & 5.16). This means that regardless of \( \alpha_T \) of the surrounding materials, this region will be highly susceptible to corrosion by ATOX though the extent of the affected region is dependent on the \( \alpha_T \) of the surrounding materials.

Table 5.4: Key ATOX flux values for a simple rectangular pit with openings of 0.1m, 0.3m and 0.5m wide and \( \alpha_T \) of 0, 0.5 and 1. For reference \( N_\infty = 4.732 \times 10^{19} \, m^{-2} \, s^{-1} \)

<table>
<thead>
<tr>
<th>( \alpha_T )</th>
<th>( l_p ) [m]</th>
<th>Max Flux ( m^{-2} , s^{-1} )</th>
<th>( N_\infty ) [m(^{-3})]</th>
<th>( N_{pit} ) [m(^{-2} , s^{-1})]</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_T )</td>
<td>( l_p ) [m]</td>
<td>FFP</td>
<td>RFP</td>
<td>FFP</td>
</tr>
<tr>
<td>0.0</td>
<td>0.1</td>
<td>3.167 \times 10^{19}</td>
<td>2.444 \times 10^{19}</td>
<td>1.773 \times 10^{18}</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>3.392 \times 10^{19}</td>
<td>2.445 \times 10^{19}</td>
<td>1.773 \times 10^{18}</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>3.437 \times 10^{19}</td>
<td>2.552 \times 10^{19}</td>
<td>1.773 \times 10^{18}</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>2.002 \times 10^{19}</td>
<td>5.385 \times 10^{18}</td>
<td>1.810 \times 10^{18}</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>2.322 \times 10^{19}</td>
<td>5.806 \times 10^{18}</td>
<td>1.811 \times 10^{18}</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>2.397 \times 10^{19}</td>
<td>6.777 \times 10^{18}</td>
<td>1.812 \times 10^{18}</td>
</tr>
<tr>
<td>1.0</td>
<td>0.1</td>
<td>1.935 \times 10^{19}</td>
<td>1.468 \times 10^{18}</td>
<td>1.848 \times 10^{18}</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>2.251 \times 10^{19}</td>
<td>1.464 \times 10^{18}</td>
<td>1.849 \times 10^{18}</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>2.321 \times 10^{19}</td>
<td>1.475 \times 10^{18}</td>
<td>1.850 \times 10^{18}</td>
</tr>
</tbody>
</table>
5.4. ATOMIC OXYGEN IN A SIMPLE PIT

Figure 5.14: ATOX flux on the Forward-facing panel (FFP) (left column) and Rearward-facing panel (RFP) (right column) of a 0.1m wide pit for $\alpha_T = 0, 0.5$ and 1. $N_\infty$ and $N_{\text{pit}}$ included in the figure for reference.
Figure 5.15: ATOX flux on the Forward-facing panel (FFP) (left column) and Rearward-facing panel (RFP) (right column) of a 0.3m wide pit for $\alpha_T = 0, 0.5$ and 1. $N_\infty$ and $N_{\rho\mu}$ included in the figure for reference.
Figure 5.16: ATOX flux on the Forward-facing panel (FFP) (left column) and Rearward-facing panel (RFP) (right column) of a 0.5m wide pit for $\alpha_T = 0, 0.5$ and 1. $N_\infty$ and $N_{\mu T}$ included in the figure for reference.
5.4.1.2 Diffuse case (α₁ = 1)

The diffuse case (α₁ = 1) was first examined in a general sense for the FFP in Section 5.3, where it was shown that the flux on this surface could be predicted using Equation 5.17. It was demonstrated in that section that the flux on the FFP is focused near the rim of the pit and that the maximum flux value is approximately half the flux experienced in the full flow. As can be seen from Figures 5.14e, 5.15e, & 5.16e and Table 5.4, this maximum flux value appears to be independent of the size of the opening or the depth of the pit, which is to be expected given the formulation of Equation 5.14. This also shows that, despite the diffuse nature of the reflections, the number density in the pit is not significantly elevated above that of the external gas and thus the flow remains broadly non-molecular in this instance.

This leads to the other observation for the FFP in Section 5.3.3, that the ambient flux in the pit (∼\(N_{\text{pit}}\)) is approximately half that of the exterior ambient flux (∼\(N_{\infty}\)) and that this too appears to be independent of the size of the pit. By contrast and as can be seen in Figures 5.14f, 5.15f, & 5.16f and from Table 5.4, the maximum flux on the RFP is less than ∼\(N_{\infty}\). However, it is slightly elevated compared to ∼\(N_{\text{pit}}\), particularly near the RFR due to the reflection and scattering of particles from the FFR.

In general, for this configuration, the ATOX ∼\(N_{\text{pit}}\) is approximately 42 times smaller than ∼\(N_{\infty}\) (or 1.9%). The ultimate implication of this is that if the pit and surrounding structures are made of materials that exhibit mostly diffuse particle-surface interactions then the region near the FFR would be the most susceptible to corrosion by ATOX. As will be seen in Section 5.4.2, the regions lower down the FFP and RFP still experience an ATOX Environment, however, the velocity and collision energy of these interactions will be much lower. This is because the particles shed most of their energy with the initial interaction with the pit near the rim.

5.4.1.3 Specular case (α₁ = 0)

In contrast to the diffuse case, the ATOX flux distribution in the specular case (α₁ = 0) is consistently high all the way down on both sides of the pit. This can be seen in Figures 5.14a, 5.15a, & 5.16a for the FFP and Figures 5.14b, 5.15b, & 5.16b for the RFP. This is to be expected since, for a specular model, the energy of the particles does not equalize with the surface and is thus re-emitted with the same energy it was absorbed, resulting in a simple flipping of the normal velocity. The particles bounce their way down and then all the way back up again, in much the same way light would within a mirrored pit. Additionally, the mean free path of the particles for the initial atmospheric conditions described in Table 5.1 is 108 metres. This is over 200x larger than the maximum size of the pits simulated, and thus the likelihood of particle-particle collisions occurring within the pit is low.

While the maximum flux on the FFP is slightly elevated compared to the diffuse case, the ‘ambient’ flux level lower in the pit is half that of ∼\(N_{\infty}\). As with the diffuse case, ‘ambient’ flux level appears to be unaffected by the size of the pit as can be seen from the Figures 5.14a, 5.15a, & 5.16a and Table 5.4. This would suggest that surfaces and equipment would be at significant risk within a pit that exhibits purely specular reflections. Since the velocity vector is simply flipped, it can be expected
that the collision energy along these walls would also be equivalent to the external maximum. It should be noted that this is an extreme example and it is not yet certain whether materials exist that exhibit specular particle-surface interactions under VLEO conditions. Further, it is known that prolonged exposure to ATOX can change the $\alpha_T$, with the particle-surface interactions becoming more diffuse as the surface slowly degrades. This has been observed from comparing predictions of satellite drag in VLEO with their actual trajectories.

5.4.1.4 Partially Diffuse case ($0 < \alpha_T < 1$)

The previous sections explored how the ATOX would behave in purely specular and diffuse situations. However, as discussed above, it is known that the $\alpha_T$ of a material can change following prolonged exposure to ATOX, with the particle-surface interactions tending to become more diffuse with time. Additionally, empirical studies have shown that the value of $\alpha_T$ in VLEO may range in value from 0.74 to 1. It is therefore important to assess what might happen in this case in these more mixed scenarios. It should be stressed that these simulations assumed that $\alpha_T$ is constant across the surfaces. In reality, there may be some uneven degradation of the materials surface as a result of the uneven distribution of ATOX fluency which could result in a varying $\alpha_T$. This feedback loop could ultimately result in a pit that has more diffuse interaction near the rim of the pit and more specular reflection on surfaces deeper down. However, the degree of degradation would be highly dependent on the material selected. This has not been explored in the work presented here and could be an avenue of further research.

Figures 5.16c & 5.16d includes an example of when the surface interactions are neither purely specular nor purely diffuse for a 0.5m wide opening, in this case when $\alpha_T = 0.5$. Similar graphs are shown in Figures 5.14c & 5.14d and Figures 5.15c & 5.15d for a 0.1m and 0.3m opening respectively. It can be seen that in addition to the peak at the top of the FFP, there are secondary and in some cases tertiary hotspots along the RFP and FFP respectively. The presence of some specular reflections ensures that these additional hotspots exist, while the presence of diffuse interactions ensure that the peak flux decays with each interaction. The more specular the reflections (i.e. the closer $\alpha_T$ is to 0) the more hotspot centres there are, until these hotspots smear out and blends together.

The lower extent of the secondary hotspot on the RFP can be predicted in much the same way as the primary hotspot described in Section 5.3.3 for the diffuse case. In a 200km orbit under the conditions in Table 5.1, the predicted $\sigma_{AD}$ for ATOX is $5.48^\circ$. For a 0.5m wide pit, this would put the lowest extent of the secondary hotspot at a depth of 0.30m which is shown as the vertical dotted line in Figure 5.16d. Unlike the diffuse case, it is not as easy to isolate the hotspots along each wall from one another or indeed the pit ambient flux. Interestingly, the peak of the secondary hotspot does appears to broadly align with $\mu$ as projected from the rim of the FFP rather than the reflection from the RFP. Projecting the hotspot through another reflection, it can be seen in Figure 5.16c that the tertiary hotspot should occur slightly higher up the pit than the drop off point of the first hotspot as predicted by $\sigma_{AD}$, which is why the flux remains high in this region.
Figure 5.17: Mean flux on the Pit floor panel (PFP) as a percentage of the external ambient flux with varying $\alpha_T$ and pit lengths of 0.1, 0.25 & 0.5m respectively.

The key implication here is that, unlike the diffuse case, there is higher flux further down in the pit and on the RFP. The surface interaction model employed does assume that some energy is lost on each interaction, so subsequent hotspots would be associated with reduced collision energy\[56]. However, as highlighted above, empirical data seems to suggest that the $\alpha_T$ of materials in VLEO is larger than 0.74 so many of these secondary hotspots will reduce significantly in intensity.

5.4.2 Depth Penetration

A critical topic to examine is how much ATOX reaches the Pit floor panel (PFP) as this is where instruments would normally be mounted. Figure 5.17 shows the mean flux experienced by the PFP as a percentage of the external ambient flux ($\tilde{N}_\infty$). As before, $\tilde{N}_\infty$ is taken to be the average flux experienced by the outer surfaces that are parallel to the bulk flow, namely those ahead and behind the pit. Each of the graphs in Figure 5.17 shows how the mean flux changes with respect to $\alpha_T$, from 0 to 1, for pits with a lengths of 0.10, 0.25 and 0.50m.

A common feature of the graphs in Figure 5.17 is that when $\alpha_T = 0$, i.e. the particle surface interaction is purely specular, the flux experienced by the PFP is equal to $\tilde{N}_\infty$. This is independent of the atmospheric conditions or dimensionality of the pit and, as was seen in the previous section, implies full depth penetration. As discussed previously, when the interactions are purely specular the particles are essentially like light in a mirrored pit. Thus in this scenario the PFP is fundamentally no different to an outer surfaces in terms of ATOX flux. This means that if the particle-surface interaction were indeed specular, there is no protective advantage to recessing sensitive equipment that needs access to the exterior of the satellite.

By contrast, when the reflections are purely diffuse ($\alpha_T = 1$), the mean flux experienced by the PFP is 40-60% of $\tilde{N}_\infty$. This is broadly in line with the $\tilde{N}_{pit}$ seen on the FFP which as was seen in Section 5.3.3 tended towards being 50% of $\tilde{N}_\infty$. This varied depending on the dimensions of the pits, with wider shallower pits experiencing the greatest penetration while deeper narrower pits experience the least. Narrower pits also saw less ATOX reaching the PFP across the energy
accommodation range, except of course when $\alpha_T$ approached 1. By contrast, there is a far smoother transition for wide shallow pits, such as the 0.50m wide by 0.1m deep pit in Figure 5.17.

What this shows is that regardless of the shape or size of the pit, it can be expected that the PFP will experience at least 40-50% of $N_\infty$. While this is still significantly lower than $N_\infty$, this may still pose a problem for delicate equipment located at or near the base. The alternative would be to simply cover the opening with a transparent material, however, this can also be susceptible to corrosion from ATOX. As this will be fully exposed to the flow this can ultimately lead to the degradation of the optical properties of the cover which in turn could also cause a reduction in the performance of that payloads or subsystems with time. If the pit surface material exhibits purely specular interactions with the ATOX then a cover may be the more appropriate route to limit the amount of ATOX that reaches the instrument. This is because there are some treatments that can be applied to the cover to offer some protection against ATOX that may not be appropriate or possible on delicate instruments[36]. If the surface interactions are more diffuse, then fewer ATOX particles can reach down into the pit. The trade-off may then be between the performance degradation of the cover versus the degradation of the instrument itself but this will be highly dependent on the payload chosen and the configuration of the satellite and pit.
5.5 Conclusion

This chapter has examined the ingress and propagation of Atomic Oxygen (ATOX) to the satellite’s internal environment. Due to its higher density in VLEO, ATOX poses a significant hazard for future missions to this region of orbit, not only for the external surfaces but also for the internal regions of the satellite.

Using both analytical methods and DSMC simulations, it was shown that, for a simple rectangular pit the maximum ATOX flux would be experienced at the FFR. If the surfaces in the pit experienced diffuse particle-surface interactions ($\alpha_T = 1$), then the maximum flux experienced at the rim was approximately half the expected flux on the front of the satellite ($N_\infty$). It was also shown that the distribution of particles on the FFP could approximate a cumulative frequency distribution of a normal distribution when $|s_\infty| > 5$. This meant that the distribution of particles could be described by the angular deviation of particles in a gas ($\sigma_{AD}$).

In cases where the surfaces were less diffuse, secondary and tertiary hotspots also formed. The lower extent of these hotspots could be predicted using the dispersion cones of $\sigma_{AD}$, while the molecular Mach angle ($\mu$) could be used to determine the centre of the hotspots. As the surface reflections became more specular, the hotspots would be blended together until there was a consistently high flux all the way down the pit on both the FFP and RFP. The mean flux along these surfaces in this case was shown to be approximately half $N_\infty$.

Finally, it was also demonstrated that as long as the iterations were at least partly diffuse ($\alpha_T > 0$) the ambient flux at the base of the pit would be approximately half the ambient flux seen on the exterior ($\bar{N}_{pit}$) or nearly 2% of $N_\infty$. While significantly smaller than $N_\infty$, this flux at the base of the pit still poses a serious risk to delicate payload and instrument. This will need to be accounted for when designing systems for VLEO.
6.1 The Challenges of VLEO

There is increasing interest in operating Earth observation and telecommunication platforms at a lower altitude, in a region known as Very Low Earth Orbit (VLEO). Despite its advantages, operating in VLEO is not without its challenges. One such challenge is combating the higher levels of atmospheric drag a satellite would experience in these lower orbits. While the drag can be compensated for using a dedicated propulsion system, the operational life will still be limited. It is therefore desirable to find ways to maximize the life of the satellite as much as possible.

In this study, Direct Simulation Monte Carlo (DSMC) was used to demonstrate that tapering the nose section of a satellite can help to reduce the drag it experiences in VLEO. This then reduces the fuel required by the platform. However, for the candidate profiles examined, this came at the expense of internal volume. It was therefore deemed beneficial to explore the interaction between the internal volume and drag of the satellite aeroshell profile.

The approach taken here was to approximate the results of the DSMC simulations using a surrogate model. By carefully selecting the sampling locations and through effective interpolation of the data, the number of simulations required to explore the design domain was significantly reduced. Thereafter, Particle Swarm Optimization (PSO) was used to optimize the satellite’s aeroshell for minimum drag and maximum volume.

Another factor considered was the higher abundance of ATOX in the VLEO environment. Since some payloads are exposed to space, they are more susceptible to the corrosive nature of ATOX which would ultimately limit their operational lives. This followed from observations of the link between the gas refill rate behind a satellite and the ideal geometry of its tail. The refill rate also had implications for the ingress and propagation of ATOX to the internal environment of the satellite. So the distribution of ATOX on the internal surfaces of a satellite was approached both analytically and with simulations, as it was felt this might help to identify relationships and simplifications that could be applied in future work.
6.2 Surrogate Model

It was noted that it would be computationally expensive to perform the simulations using Direct Simulation Monte Carlo (DSMC) methods, because of the convergence time of the particle stream. Despite this, it was still deemed beneficial to use DSMC methods for examining the geometry. Unlike faster methods, such as the panel methods developed by Sentman [54], DSMC methods can account for all aspects of the geometric effects on the flow, including any non-linearities. This includes reflections and the increased particle-particle encounters in the higher density regions around the satellite. However, the long simulation time was a major concern and would normally make it prohibitive to explore an entire design space with sufficient fidelity.

The computational complexity was effectively managed by employing a surrogate model framework that reduced the number of simulations required by carefully selecting the sample points and interpolating between them. Sampling of the design space was achieved using a multi-criteria adaptive sampling framework based on the work of Mackman et. al. [69]. The criterion they devised is the product of the Laplacian and the separation function and thus gives a balance between adding points in locations of high detail and unsampled regions of the domain, which ensured a degree of space-filling to minimize uncertainty in less detailed regions.

For the 6 parameter Blunted Wedge profile analysed in Chapter 4, the maximum time spent evaluating the adaptive criteria was 1 minute and 5 seconds per sample (17 minutes and 24 seconds for a batch of 16). While this duration is not negligible, it is still significantly shorter than the 1 hour 42 minutes required to fully simulate the body and also occurred during the last adaptive evaluation which had the largest matrix to invert. Therefore, the generation process remained dominated by the samples themselves and not the process to find them.

Table 6.1 provides a comparison the performance of the generator using adaptive sampling against a simpler sample scheme such as Fully Factorial (FF) for the Blunted Wedge profile with 2, 4, and 6 active design parameters under similar atmospheric conditions (Table 4.1). As can be seen from Table 6.1 a gridded approach to the Blunted Wedge with 6 parameters would have required at least 4,826,809 samples to achieve the same level of uncertainty in the model as the adaptive approach. This is 707× more samples than were needed under the adaptive sampling used here. Without the reduction, the DSMC simulations alone would take about 44 years using 16 cores. While this run time could be improved by running the DSMC simulations over more cores, approximately 10725 cores would be needed to complete the simulations in the same time frame as the adaptive sampling method.

<table>
<thead>
<tr>
<th>n</th>
<th>Adaptive</th>
<th>Gridded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>√N</td>
</tr>
<tr>
<td>2</td>
<td>177</td>
<td>13</td>
</tr>
<tr>
<td>4</td>
<td>2100</td>
<td>6.76</td>
</tr>
<tr>
<td>6</td>
<td>6827</td>
<td>4.36</td>
</tr>
</tbody>
</table>
Upon being deployed for the work in Chapter 4, it was noted that there were specific instances where the results from the surrogate models were less precise (as discussed in Section 4.3). This resulted in the formation of local minima in certain areas of the design space. However, these local minima were usually in regions where the combination of parameters would have had little or no effect on the output value (such as the fixed $r_t$ case for the Blunted Wedge in Section 4.3.3) and these local minima could be handled by the Particle Swarm Optimization (PSO).

While the surrogate model did present a number of challenges during its development, as can be seen from the summary analysis above, it proved to be a suitable approach for this work. The surrogate model generator provided a framework through which the parameter space of the satellite geometry could be explored using DSMC simulations and as can be seen from the results presented in Chapter 4 this was successfully achieved for several aeroshell profiles in Very Low Earth Orbit (VLEO) under a variety of scenarios.

### 6.3 Shape Optimization

Using the surrogate model generator developed in Chapter 3 together with DSMC simulations, the effectiveness of altering the geometry of the satellite to reduce its drag was examined. The profiles examined included a Blunted Wedge, Elliptical, Double-Conic nose and Rounded-Conic profiles (see Figure 6.1).

![Example candidate profiles](image_url)

(a) Blunted Wedge  
(b) Elliptical  
(c) Double-Conic  
(d) Rounded-Conic

Figure 6.1: Example candidate profiles to be examined in this section
Figure 6.2: Graphs showing the Pareto-optimal fronts when optimizing for minimum drag and maximum volume
While performing initial explorations of the atmospheric drag experienced by satellites in VLEO, it was demonstrated that a tapered nose section can help to reduce the drag and, in turn, reduce the fuel required by the platform. Using the DSMC surrogate model, it has been shown that a maximum drag reduction of between 21% to 35% was achievable when the aspect ratio of the profile was fixed at 1:2 and 1:6 respectively. This was achieved using a simple conic nose profile that occupied the full length of the central body. While these were significant reductions in the drag, for bodies with fixed aspect ratios this also resulted in a significant loss of internal volume when compared to a similarly sized cuboid. This made it interesting to explore the interaction between volume and drag for the aeroshell profile.

Using multi-objective optimization, a set of Pareto-optimal satellite geometries were identified that traded-off minimizing the drag against maximizing the internal volume of the profiles. The Pareto-optimal fronts for the Blunted Wedge and Elliptical profiles are shown in Figures 6.2a & 6.2b respectively. When the aspect ratio is fixed, the Elliptical profile generally performed the best of the 4 profiles tested, particularly for higher aspect ratio bodies. This was especially the case when considering the trade-off between drag and fuel volume.

Because the fuel mass and its volume is proportional to the drag, a reduction in the drag will result in a reduction in required fuel system volume. The reduction in the required fuel system volume (or the drag) should be equal to or less than the loss of internal volume due to the aeroshell. It was seen that for bodies with an aspect ratio of 1:6, the Elliptical profile achieved a drag reduction of 26% with a reduction in the volume of just 22%. By comparison, the Blunted Wedge profile achieved a reduction in both the drag and volume of 22%.

Figure 6.3: Improvement in atmospheric drag using a Blunted Wedge profile compared to a Cuboid profile of similar volume when either the length or the width of the body is fixed.
It was also generally observed that the 'nose-only' profiles did not perform as well as either the Blunted Wedge or Elliptical profiles, even when the optimizer was solely optimizing for drag. The key difference was the lack of a tail profile on these morphologies. In particular, it was observed that beyond a certain internal tail angle, the tapered surface was effectively shielded by the body of the satellite and thus had a limited contribution to the overall drag experienced. This ultimately proved beneficial while optimizing the satellite geometry, especially for those geometries with high aspect ratios.

While the external dimensions can often be a constraining factor on the geometry of a satellite, it is more common to prescribe a desired internal volume for the satellite body. Figure 6.3 shows the possible improvements in the drag when the volume is prescribed using a Blunted Wedge profile. To help explore the design space, the length and width of the body were constrained separately, to simulate the dimensional needs of internal subsystems (Figures 6.3a & 6.3b respectively). Figure 6.4 provides an example of optimized aeroshell geometries for a 0.5m$^2$ body using a Blunted Wedge profile when the length and width are constrained. As can be seen from Figure 6.3a, it was shown that when the length of the body was constrained, only a marginal decrease in the drag could be achieved using any of the profiles. This was because the body has to increase its frontal area to accommodate the profile (see Figure 6.4a).

By contrast, when the width was constrained, there was little variation in the drag experienced by each of the profiles as the volume increased. This was possible because the bodies were able to grow in length to accommodate the profiles (see Figure 6.4b), thus achieving a far greater reduction in drag. It was also demonstrated that simply attaching a tapered profile to the front of the satellite body would provide a reduction in drag.

The effectiveness of altering the geometry of the satellite to reduce its drag was assessed given a number of constraints and scenarios. The reduction in drag was traded-off against the loss of internal volume to find practical aeroshell profiles for specific volumes. While there may be other constraints on the design which might drive the selection, the families of aeroshell profiles presented in this work will provide useful guidance for future VLEO platform design.

(a) Fixed Length = 1.0m
Drag = 13.41mN (Improv.= 2.81%)

(b) Fixed Width = 0.5m
Drag = 11.87mN (Improv.= 13.97%)

Figure 6.4: Optimized aeroshell geometries for a 0.5m$^2$ body using a Blunted Wedge profile
Original Cuboid drag = 13.80mN
6.4 Atomic Oxygen

Most prior research into the effects of Atomic Oxygen (ATOX) has focused on its impact on the external portions of the satellite. However, due to the higher abundance of ATOX in the VLEO region, other areas of the satellite are at a greater risk of exposure, critically, this includes payloads and other subsystems that require access to space. This followed from observations of the link between the gas refill rate behind the satellite and the ideal geometry of the tail. The refill rate also has implications for other aspects of the geometry, in particular the ingress and propagation of ATOX to the internal environment of the satellite. The problem was approached both analytically and with simulations, as it was felt this might help to identify relationships and simplifications that could be applied in future work.

The analytical approach provided an interesting insight into how gases disperse in the rarefied environment of VLEO. Derivations were made for both the angular flux distribution and the flux distribution along a wall in Chapter 5 (full derivations in Appendix C). It was found that the necessary integration was not straightforward and many simplifications were necessary to obtain the final form. For this, it was useful to have DSMC results to compare the predicted distributions. These results helped to verify the derivation and ultimately it was seen in Section 5.3.3 that good agreement was achieved between the derived analytical model and the DSMC results when particle-surface interactions were diffuse.

The analytical results allowed for some useful properties to be extracted and applied to the other types of gas-surface interactions that were simulated using DSMC. For instance, it was shown that in a simple rectangular pit, the highest ATOX flux would be experienced at the Forward-facing rim (FFR) (see Figure 6.6). If the surfaces in the pit experienced diffuse particle-surface interactions ($\alpha_T = 1$), then the maximum flux experienced at FFR was approximately half the expected flux.

![Figure 6.5: Diagram showing the velocity vectors and the projected angular probability distribution function $F(\theta)$](image)
Figure 6.6: Diagram showing the simulation set up for a simple rectangular pit, including key dimensions and the locations of the Forward-facing panel (FFP), Rearward-facing panel (RFP), Pit floor panel (PFP), Forward-facing rim (FFR) and Rearward-facing rim (RFR) on the front of the satellite ($N_\infty$). It was also shown that the distribution of particles on the Forward-facing panel (FFP) could approximate a cumulative frequency distribution of a normal distribution when the $|s_\infty| > 5$. This meant that the distribution of inbound particles could be described by the angular deviation of particles in a gas ($\sigma_{AD}$).

While the analytical results helped understand the simulated data, the models had some limitations. For instance, the model did not account for reflections and thus could only predict the fluency of incoming particles from the far-field. Since the particle-surface interactions affect how the gases propagate in the cavity, it would be useful to be able to predict analytically how the ATOX fluency on the surfaces changes with the $\alpha_T$. However, these inbound ATOX particles also have the highest energy and would therefore cause the most damage to the surfaces of the satellite. Therefore, knowing which surfaces these inbound particles will affect and to what extent is still important for material selection and risk mitigation.

To examine how the ATOX particles propagate within a cavity, a set of DSMC simulations were run on a simple rectangular pit. These simulations were performed across a range of energy accommodation coefficients, as there is some uncertainty around the value of $\alpha_T$ for materials in VLEO. Empirical findings from satellites in orbit [27, 32] seem to conflict with ground-based tests of materials.

Regardless of the surface properties or dimensions of the pits, it was observed that the most susceptible region of the pit was the FFR at the top of the FFP. The flux here was always at least half $N_\infty$ but rose to 75% $N_\infty$ when $\alpha_T = 0$. For the diffuse cases ($\alpha_T = 1$), the susceptible regions of the FFP formed a hotspot near the FFR. The lowest extent of this hotspot could be predicted using the dispersion cones of $\sigma_{AD}$ as shown in Figure 6.7.
It was also observed that in cases where the surfaces had less diffuse interactions, secondary and tertiary hotspots also formed down the surfaces of the FFP and Rearward-facing panel (RFP). The lower extent of these hotspots could be predicted using $\sigma_{AD}$ by reflecting the dispersion cones. As the surface reflections became more specular, the hotspots would be blended together until there was a consistently high flux all the way down the pit on both the FFP and RFP. The mean flux along these surfaces when $\alpha_T = 0$ was shown to be approximately half $\tilde{N}_\infty$.

In general, it was demonstrated that regardless of the size of the pit, a surface can expect to experience at least 40-50% of $\tilde{N}_\infty$ or nearly 2% of $N_\infty$. While this is still significantly lower than $N_\infty$, this may still pose a problem for delicate equipment located at or near the base of the pit. If the pit surface material exhibits purely specular interactions with ATOX, then a cover may be useful to limit the amount of ATOX that reaches the instrument. This is because there are some treatments that can be applied to a cover to offer some protection against ATOX that may not be appropriate or possible on delicate instruments\cite{36}. While the exact nature of the particle-surface interactions under VLEO conditions is not fully understood\cite{22}, it is known that prolonged exposure to ATOX can change the $\alpha_T$, with the particle-surface interactions becoming more diffuse as the surface slowly degrades.

If the surface interactions are more diffuse, then fewer ATOX particles can reach down into the pit, but still at least 40% of $\tilde{N}_\infty$. The trade-off may then be between the performance degradation of the cover versus the degradation of the instrument itself. This will be highly dependent on the payload chosen and the configuration of the satellite and pit.

Figure 6.7: Distribution of Atomic Oxygen ($O_1$) flux on the FFP of a recessed pit at 200km in altitude

![Graph showing the distribution of atomic oxygen flux](image)
6.5 Future Research Topics

Over the course of the work performed as part of this thesis, a number of topics have been identified for future research.

For instance, the use of a panel method (such as the one developed by Sentman) as the underlying trend model in a surrogate model could help to improve the accuracy in the test case used here. It was demonstrated that including a polynomial function provides no benefit, however, it is possible that the ‘black box’ functions tested were not suitable. A panel method may capture the trend of the data far better which might ultimately reduce the number of samples needed by the surrogate model.

It was also seen that switching to the Power Function to calculate the uncertainty of the surrogate model significantly improved the speed of the adaptive process. However, this process is still relatively computationally complex and for larger runs it still made sense to limit the frequency with which the uncertainty was calculated. It would be useful to identify or develop faster methods for determining the uncertainty in the model so that this can be tracked fully during the generation process.

To enable the investigation into the effectiveness of altering the geometry of the satellite to reduce the drag, as presented in Chapter 4, some assumptions had to be made about the design space and in particular the geometries of the satellite. For instance, given the uncertainty surrounding the gas surface interactions in VLEO, the energy accommodation coefficient was set as 0.95 indicating a mostly diffuse interaction. Future analysis could determine whether the energy accommodation coefficient has any significant impact on the choice of profile.

Additionally, the profiles examined in Chapter 4 were all two dimensional. It would be useful to extend the analysis to three-dimensional geometries to see if the relationships identified in this research still hold. Extending into three dimensions would also allow the consideration of externally mounted subsystems, such as the solar panels, since these were not considered in this analysis. The solar panels often represent a significant contribution to the drag a satellite experiences, so incorporating them represent a vital link in the drag-power optimization loop. Careful consideration would be needed into how they are mounted to the body as this will affect their contribution to the drag. This could then lead to a broader systems analysis with the geometry of the satellite as part of the loop.

The work performed in Chapter 5 focused purely on the influx of gases into a set of simple rectangular pits. In reality, the interior of a satellite and its interface with the outside world can be far more complex. Future work could examine the distribution of atomic oxygen when the cavity has overhanging or branching sections. Additionally, cavities with multiple openings, such as on the opposite side of the satellite, may also be of interest.

One key element that was not explored as part of this is the change in flux with respect to the angle of attack. Changing the orientation of the pit affects how shielded the internal region is and thus the distribution of ATOX that can be expected.
6.6 Closing Remarks

Using multi-objective optimization a set of Pareto-optimal satellite geometries were identified that traded-off minimizing the drag against maximizing the internal volume of the profiles. When the aspect ratio of the body is fixed, it was observed that the Elliptical profile generally performed the best of the 4 profiles tested, particularly for higher aspect ratio bodies. This is because the Elliptical profiles provided a good compromise between low drag and large volume. However, for a given internal volume it was shown that a Blunted Wedge profile is able to achieve a greater reduction in the drag experienced when compared to a similarly sized cuboid body. It was also observed that, in general, increasing the length of the satellite body to incorporate an aerodynamic profile is always beneficial.

The higher abundance of ATOX in the VLEO environment and its impact on the internal environment of the satellite was also examined. This problem was approached both analytically and with simulations as this identified relationships and simplifications that could be applied in future work. It was shown that the maximum ATOX flux inside a simple rectangular pit was experienced at the rim of the forward-facing panel and is approximately half the flux seen on the front of the satellite in the direct flow. It was also observed that the angular distribution of particle through a point in space approximates a normal distribution in the orbital environment. This also means that the flux distribution on the forward-facing surface can be approximated by the cumulative frequency distribution of this normal distribution.

Operating in VLEO presents a number of challenges including the higher level of atmospheric drag a satellite would experience and the higher abundance of ATOX in this environment. The insights gained into optimized aero-geometry and the internal flow dynamic advance the understanding both from an academic and a commercial industrial perspective. The recommendations presented here will help to inform the design of future VLEO missions and the modelling tools developed can be applied in their current form or extended to support further research into this field.
A.1 A review of the precision and accuracy of the DSMC software SPARTA

A.1.1 Introduction

In Chapter 2, DSMC was chosen as the primary method for analysing the aerodynamic forces in this thesis. While panel methods and Ray-Tracing Panel (RTP) are useful for establishing an estimate of the drag, it is desirable to account for all aspects of the geometric effects including any non-linearities. This includes reflections and the increased particle-particle encounters in the higher density regions around the satellite (of which the latter is not accounted for in Test-Particle Monte Carlo (TPMC)).

An important consideration is the time taken to perform the simulations. Panel methods, such as the Sentman method, are generally very quick as the computational complexity normally scales with the number of surfaces of the geometry. The candidate profiles used in Chapter 2 have a maximum of 6 surfaces, so the simulation time was less than a second per profile. DSMC simulations are computationally more complex as they simulate the particle physics of the rarefied gas around the satellite with the complexity scaling with the number of particles simulated. However, some DSMC implementations, like Stochastic Parallel Rarefied-gas Time-accurate Analyzer (SPARTA) use a reduced number of simulated particles to represent the molecules in the gas. In this instance the simulated particles number $f_{num}$ is the number of real particles that a single simulated particle represents in SPARTA. This then relates the true number density $N_D$ to the simulated number density $N_S$ such that

$$N_S = \frac{N_D}{f_{num}}. \quad (A.1)$$

The run time of the simulation is directly proportional to the number of particles that are simulated, with more particles requiring more time. It is important that sufficient particles are simulated to achieve a reliable result, however, this must be weighed up against the computational time this
requires. The aim of this section is to identify the minimum number of time steps and minimum $N_S$ required to produce a valid and accurate result. Under ideal circumstances a reliable result should have an uncertainty in the $C_D$ of less than $\pm 1 \times 10^{-3}$, i.e. an error of $\pm 0.1\%$. This is important as the variation in the $C_D$ is expected to be no larger than $10\%$.

### A.1.2 Method

To perform the analysis a series of simulations were required. These used the two-dimensional 0.5mx1m S1 body defined in Section 2.2.2 in a 200km orbit (see Table 2.3 for the atmospheric conditions). The $N_S$ was varied from $1 \times 10^0$ to $1 \times 10^6$ in step of 0.1 in the order. At each $N_S$, the simulation was run 10 times with a different seed to help randomize the results and capture the convergence trend.

Each simulation was run for 10000 time steps. Every 100 time steps the forces on the simulated body were calculated using a running average and the previous 100 time steps.

### A.1.3 Results & Discussion

Figure A.1a shows how the $C_D$ of the S1 test body converges as the number of time steps is increased. Figure A.1b shows the error in the $C_D$ compared to the final value achieved by that simulation as the time steps increase. Included on both figures are simulations with $N_S$ of between $1 \times 10^1$ to $1 \times 10^5$.

As can be seen from Figure A.1, at the start of a simulation there is a large error in the drag value but this quickly converges to near the final value. Within 3000 time steps most of the simulations have an error of less than $1 \times 10^{-2}$ as compared to the final value after 10000 time steps. This is consistently observed for all particle densities higher than $1 \times 10^4$ particle density. As the simulation continues to run, the result slowly converge towards the final value.

What this suggests is that for a quick results in the right ball park, the SPARTA simulations should be run for at least 2000 to 3000 time step. However, to achieve an uncertainty in the result of less than $1 \times 10^{-3}$, the simulation must be run for at least 8000 time steps. In general the simulation time is proportional to the number of time steps so to achieve this greater level of uncertainty the simulation would be run for 3-4 times longer.

In general, the higher the $N_S$, the less uncertainty there is in the result. This can be seen in Figure A.2 which shows how the uncertainty for SPARTA simulations of the S1 test body decreases as $N_S$ is increased. This would suggest using a value for the $N_S$ that is as high as possible (or an $f_{num}$ that is as small as possible) thus fewer real particles are being represented by one simulated particle.

However, increasing the number of simulated particles also increases the simulation time as can be seen by Figure A.3. As can be seen from Figure A.2, to achieve an uncertainty in the $C_D$ of less than $\pm 1 \times 10^{-3}$ the $N_S$ should be at least $1 \times 10^4$. For 10000 time steps, this has a run time of $5.5 \times 10^8$ s (or 1.5hrs) on a single core.
A.1. A REVIEW OF THE PRECISION AND ACCURACY OF THE DSMC SOFTWARE SPARTA

Figure A.1: Convergence of a set of similar SPARTA simulations on the S1 test body as the number of time steps is increased

A.1.4 Conclusion

The speed of the simulation is directly proportional to the number of particles that are simulated, with more particles requiring more time. Under ideal circumstances a reliable result should have an uncertainty in the $C_D$ of less than $\pm 1 \times 10^{-3}$. To achieve this for two-dimensional simulations in SPARTA, the simulated particle density $N_S$ should be at least $1 \times 10^4$. Assuming the simulation runs for 10000 time steps, this would have a run time of 1.5hrs on a single core.
Figure A.2: Convergence of a set of similar SPARTA simulations on the S1 test body as the number of simulated particles is increased.
Figure A.3: Relationship between the number of simulated particles and the run time

Figure A.4: Error in the $C_D$ against the simulation run time
B.1 Inverting the system matrix of a Surrogate Model

As described in Section 3.3.1.1, a Radial Basis Function (RBF) surrogate model can be expressed in matrix form as

\[
\begin{bmatrix}
\mathbf{y} \\
\mathbf{0}
\end{bmatrix} = \begin{bmatrix}
\mathbf{R} & \mathbf{F} \\
\mathbf{F}^T & 0
\end{bmatrix} \begin{bmatrix}
\beta \\
\gamma
\end{bmatrix}
\]  

\( (B.1) \)

where

\[
\mathbf{R} = \begin{bmatrix}
\phi_{1,1} & \phi_{1,2} & \ldots & \phi_{1,N} \\
\phi_{2,1} & \phi_{2,2} & \ldots & \phi_{2,N} \\
\vdots & \vdots & \ddots & \vdots \\
\phi_{N,1} & \phi_{N,2} & \ldots & \phi_{N,N}
\end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix}
f_{1,1} & \ldots & f_{1,M} \\
f_{2,1} & \ldots & f_{1,M} \\
\vdots & \ddots & \vdots \\
f_{N,1} & \ldots & f_{N,M}
\end{bmatrix}
\]  

\( (B.2) \)

\( \mathbf{R} \) contains all pairwise combinations for the basis function such that \( \phi_{i,j} = \phi(||\mathbf{x}_i - \mathbf{x}_j||) \) where \( \mathbf{x}_i \) and \( \mathbf{x}_j \) are both points in the training data \( \mathbf{X} \). \( \mathbf{F} \) contains the monomial results for each sample point \( \mathbf{x}_i \) such that \( f_{i,k} = f_k(\mathbf{x}_i) \). \( \beta \) and \( \gamma \) are the column vectors of coefficients for the basis function and polynomial respectively.

The coefficients of the surrogate model, \( \beta \) and \( \gamma \), can be obtained by inverting the systems matrix as shown in Equation \( [B.3] \). However, as a result of the large zero block in the bottom right of the matrix, the system matrix can be poorly conditioned and runs the risk of being singular to the point of precision when inverted (see Appendix \( [B.3] \) for more information). It is therefore useful to find equations for the solutions to \( \beta \) and \( \gamma \) that do not explicitly require the inversion of the full systems matrix.

\[
\begin{bmatrix}
\beta \\
\gamma
\end{bmatrix} = \begin{bmatrix}
\mathbf{R} & \mathbf{F} \\
\mathbf{F}^T & 0
\end{bmatrix}^{-1} \begin{bmatrix}
\mathbf{y} \\
\mathbf{0}
\end{bmatrix}
\]  

\( (B.3) \)
The systems matrix in Equation B.1 is a block matrix, with sub-matrices $R$, $F$, $0$. Since $R$ and $O$ are square matrices, the system matrix can be inverted blockwise such that

$$
\begin{bmatrix}
    R & F \\
    F^T & 0
\end{bmatrix}^{-1} = 
\begin{bmatrix}
    R^{-1} - R^{-1}F(F^TR^{-1}F)^{-1}F^TR^{-1} & R^{-1}F(F^TR^{-1}F)^{-1} \\
    (F^TR^{-1}F)^{-1}F^TR^{-1} & -(F^TR^{-1}F)^{-1}
\end{bmatrix} 
$$

(B.4)

By inserting the blockwise inverted systems matrix, Equation B.3 becomes

$$
\begin{bmatrix}
    \beta \\
    \gamma
\end{bmatrix} = 
\begin{bmatrix}
    R^{-1} - R^{-1}F(F^TR^{-1}F)^{-1}F^TR^{-1} & R^{-1}F(F^TR^{-1}F)^{-1} \\
    (F^TR^{-1}F)^{-1}F^TR^{-1} & -(F^TR^{-1}F)^{-1}
\end{bmatrix}
\begin{bmatrix}
    y \\
    0
\end{bmatrix} 
$$

(B.5)

By partitioning Equation B.5, $\beta$ can be rewritten in terms of Equation B.6 and $\gamma$ can be rewritten in terms of Equation B.7

$$
\beta = (R^{-1} - R^{-1}F(F^TR^{-1}F)^{-1}F^TR^{-1})y 
$$

(B.6)

$$
\gamma = (F^TR^{-1}F)^{-1}F^TR^{-1}y 
$$

(B.7)

Given Equation B.7, $\beta$ can also be rewritten in terms of $gamma$ and $y$

$$
\beta = R^{-1}(y - F\gamma) 
$$

(B.8)
B.2 Comparison of Sampling Strategies

B.2.1 Introduction

An important aspect of the surrogate model generation is the choice of sampling strategy. As was seen in Section 3.3.2, these vary from simple single-stage sampling methods, such as the Fully Factorial and Optimized Latin Hypercube, to more complex methods that adapt to the function being examined. From the literature[59], adaptive sampling schemes were identified as a preferred sampling strategy for aerodynamic research. This section of work will assess whether adaptive methods are appropriate for the work being done as part of this thesis.

B.2.2 Method

In this section four sampling methods where considered: the two single-stage sampling methods: Fully Factorial (FF) (Section 3.3.2.2) and Optimized Latin Hypercube Sampling (OLHS) (Section 3.3.2.3) and the Adaptive strategy (Section 3.3.2.4) using either FF (FF+Adapt) or OLHS (OLHS+Adapt) as the first stage. These were used to generate surrogate models for the four test cases described in Section 3.3.3.1, Droplet Function, Franke Function, Aero-1 and Aero-2 using a range of total samples. To allow direct comparison between Fully Factorial and the other sampling methods only sample budgets that were a squared number were considered. Similarly, to compare between the initial Latin Hypercube and Fully Factorial for the adaptive methods, the initial number of samples was taken to be one-third of the total number of samples[59] and then rounded to the nearest squared number (for instance, with a sample budget of 100 the initial sample budget would be 36 rather than 33). The resulting models were compared against a 51x51 grid of each test case as described in Section 3.3.4.1.

B.2.3 Result & Discussion

The results of the test scenarios are presented in Figures B.1 & B.2 with a detailed breakdown of the results for a sample budget of 100 points in Table B.1. The Root Mean Squared Error (RMSE) and maximum error in Figures B.1 & B.2 respectively were calculated using the Gridded Evaluation method described in Section 3.3.4.1 with a 51 x 51 evaluation grid. Figure B.3 provides examples of the surrogate models that can be generated for the Droplet Function with 100 sample for the four strategies (FF) Figure B.3a, FF + Adapt Figure B.3b, OLHS Figure B.3c, OLHS + Adapt Figure B.3d as well as the absolute error in these models (FF) Figure B.3e, FF + Adapt Figure B.3f, OLHS Figure B.3g, OLHS + Adapt Figure B.3h.

It is important to highlight that the fluctuation in the FF sampling methods for the Droplet Function (Figures B.1a & B.2a) are mostly due to whether a sample happens to be coincident with the central peak for a given sample budget. In all other cases, the FF has a strong negative correlation with the sample budget. In general OLHS also follows a similar trend to the FF but fluctuates around the gridded sampling method, which is likely due to the stochastic nature of the OLHS method. For a given sample budget, the adaptive sampling, as applied to the test cases, generally achieved the lowest overall RMSE and maximum error. This is as expected given the literature and therefore provides confirmation as to the validity of the approach, especially for the
APPENDIX B. NOTES ON THE DEVELOPMENT OF THE SURROGATE MODEL

(a) Droplet Function

(b) Franke’s function

(c) AERO-1: Varying length and height

(d) AERO-2: Fixed dimensions and varying nose geometry

Figure B.1: Comparison of RMSE for each sampling strategies with varying number of samples and fixed support radius $R_s = \sqrt{2}$

two aerodynamic test cases: Aero-1 (Figures B.1c & B.2c) and Aero-2 (Figures B.1d & B.2d). For the adaptive methods, there was no significant difference between using either FF or OLHS as the initial stage.

The only major exception was Franke’s function, specifically the FF sampling method achieved an RMSE similar to the two adaptive sampling methods. However, this was not observed for the maximum error and, in general, the adaptive methods did perform better than the OLHS method. The primary reason for the similarity in the RMSE could be the nature of the function itself. Franke’s function and the Droplet function were chosen as they provide competing requirements for the adaptive criteria. The droplet function, for instance, has a region of significant curvature concentrated in the centre of the domain. Under the adaptive scheme, the selection of new samples is dominated by the Laplacian component of the adaptive criteria and mediated by the separation function to ensure good coverage. As can be seen from Figure B.1a, this results in a significant improvement in the uncertainty over the single-stage methods.
B.2. COMPARISON OF SAMPLING STRATEGIES

(a) Droplet Function

(b) Franke’s function

(c) AERO-1: Varying length and height

(d) AERO-2: Fixed dimensions and varying nose geometry

Figure B.2: Maximum observed error for each sampling strategies with varying number of samples and fixed support radius $R_s = \sqrt{2}$

Franke’s function, by contrast, has more peaks but these are much smoother and thus the curvature is much lower and far more spread throughout the domain. This lends itself more to the space-filling aspect of the adaptive criteria, and in that sense, both the adaptive methods and the FF method achieve a consistent coverage of the domain. However, the Laplacian component still provides a bias towards the curvier region of the design space (such as the peaks) which helped to reduce the maximum error. This is an encouraging outcome as it shows that the adaptive methods, as coded here, can be applied to a range of functions and achieve a level of uncertainty that is as good as or better than the single-stage methods.

A key argument for using adaptive sampling is to achieve improvements in both the generation speed and the accuracy of the surrogate model. However, despite the improvements in accuracy, for the analytical functions examined here, the adaptive method is much slower. For instance, for a given sample budget the adaptive sampling takes between 15 and 30 times longer to run than
APPENDIX B. NOTES ON THE DEVELOPMENT OF THE SURROGATE MODEL

Figure B.3: Generated surrogate model and absolute error in model compared to true function of the Droplet function for the 4 sampling strategies for $N = 100$
Table B.1: A comparison of sampling strategies with maximum sampling points of 100

<table>
<thead>
<tr>
<th>Test Case</th>
<th>Sampling</th>
<th>min value</th>
<th>max value</th>
<th>RMSE Error</th>
<th>Max Error</th>
<th>Run time [s]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Droplet</td>
<td>FF</td>
<td>-2.637</td>
<td>0.60</td>
<td>2.07 × 10⁻¹</td>
<td>2.40 × 10⁰</td>
<td>2.17</td>
</tr>
<tr>
<td></td>
<td>OLHS</td>
<td>-2.746</td>
<td>3.00</td>
<td>1.26 × 10⁻¹</td>
<td>1.05 × 10⁰</td>
<td>2.31</td>
</tr>
<tr>
<td></td>
<td>FF + Adapt</td>
<td>-2.680</td>
<td>3.00</td>
<td>2.15 × 10⁻²</td>
<td>1.48 × 10⁻¹</td>
<td>59.16</td>
</tr>
<tr>
<td></td>
<td>OLHS + Adapt</td>
<td>-2.683</td>
<td>3.00</td>
<td>2.13 × 10⁻²</td>
<td>1.77 × 10⁻¹</td>
<td>59.69</td>
</tr>
<tr>
<td>Franke</td>
<td>FF</td>
<td>-0.197</td>
<td>1.037</td>
<td>1.23 × 10⁻³</td>
<td>9.73 × 10⁻³</td>
<td>2.05</td>
</tr>
<tr>
<td></td>
<td>OLHS</td>
<td>-0.187</td>
<td>1.039</td>
<td>6.02 × 10⁻⁰</td>
<td>4.25 × 10⁻²</td>
<td>2.12</td>
</tr>
<tr>
<td></td>
<td>FF + Adapt</td>
<td>-0.190</td>
<td>1.039</td>
<td>3.56 × 10⁻³</td>
<td>2.06 × 10⁻²</td>
<td>59.31</td>
</tr>
<tr>
<td></td>
<td>OLHS + Adapt</td>
<td>-0.188</td>
<td>1.039</td>
<td>2.70 × 10⁻³</td>
<td>1.36 × 10⁻²</td>
<td>59.26</td>
</tr>
<tr>
<td>Aero-1</td>
<td>FF</td>
<td>-8.485</td>
<td>-2.118</td>
<td>2.60 × 10⁻¹</td>
<td>2.02 × 10⁰</td>
<td>7.10</td>
</tr>
<tr>
<td></td>
<td>OLHS</td>
<td>-8.485</td>
<td>-1.826</td>
<td>3.13 × 10⁻¹</td>
<td>3.05 × 10⁰</td>
<td>6.67</td>
</tr>
<tr>
<td></td>
<td>FF + Adapt</td>
<td>-8.485</td>
<td>-2.107</td>
<td>9.24 × 10⁻²</td>
<td>1.19 × 10⁰</td>
<td>118.41</td>
</tr>
<tr>
<td></td>
<td>OLHS + Adapt</td>
<td>-8.485</td>
<td>-2.096</td>
<td>1.14 × 10⁻¹</td>
<td>1.04 × 10⁰</td>
<td>110.83</td>
</tr>
<tr>
<td>Aero-2</td>
<td>FF</td>
<td>-2.384</td>
<td>-2.192</td>
<td>1.86 × 10⁻³</td>
<td>1.17 × 10⁻²</td>
<td>7.32</td>
</tr>
<tr>
<td></td>
<td>OLHS</td>
<td>-2.384</td>
<td>-2.192</td>
<td>3.08 × 10⁻³</td>
<td>2.49 × 10⁻²</td>
<td>7.27</td>
</tr>
<tr>
<td></td>
<td>FF + Adapt</td>
<td>-2.385</td>
<td>-2.192</td>
<td>9.08 × 10⁻⁴</td>
<td>8.75 × 10⁻³</td>
<td>115.28</td>
</tr>
<tr>
<td></td>
<td>OLHS + Adapt</td>
<td>-2.385</td>
<td>-2.192</td>
<td>8.34 × 10⁻⁴</td>
<td>6.86 × 10⁻³</td>
<td>116.77</td>
</tr>
</tbody>
</table>

The equivalent single-stage sampling method as can be seen in Table B.1. The longer generation time is mostly a result of calculating the criteria and then finding the new sample. The run time on the equations and the panel methods is fixed across methods and relatively small. One of the advantages of adaptive sampling is that for a given uncertainty fewer samples are needed, as can be seen in Figure B.1. However, as is seen in Figure B.4, the run time is always longer using an adaptive method. Therefore, for the analytical functions used here, including the panel methods, single-stage methods are sufficient as they are far faster at generating a surrogate model for a given uncertainty than an adaptive scheme that might use fewer samples.

Of course, the primary reason for investigating surrogate models in the first place is to use them in conjunction with DSMC simulations which take significantly longer to run per sample than panel methods. For instance, the panel method used for the aerodynamic test cases took on average 0.06 seconds to generate the result per sample. By contrast, the simulations time required for a comparable sample can take several orders of magnitude longer to run using DSMC methods. Since the ultimate aim is to use DSMC methods on the satellite bodies to determine the aerodynamic forces, it is important to consider what impact this will have on the total run time of the surrogate model.

Figure B.5 shows the RMSE plotted against the generation time assuming a DSMC simulation time of 1 hour per sample (see section Section 2.5.3) for the Aerodynamic test cases. Parallel computation has not been accounted for in this case but would result in the same speed up and overheads for all methods. As can be seen from Figure B.5, the adaptive methods are faster for a given uncertainty and more accurate for a given run time assuming a DSMC simulation time. In general, it was observed that so long as the time taken to determine a new sample point was much smaller than the time taken to collect the sample, the adaptive methods were always better for a given surrogate model uncertainty.
The effectiveness of the adaptive sampling method appeared to be largely independent of the initial sampling method employed. For instance, as can be seen from Table B.1, there is no significant difference between either the RMSE or Max Error of the two starting sample for any of the test cases. Nevertheless, of the initial sample strategies, OLHS will be carried forward as the basis for the generation process. Arguably, the OLHS method provides a greater opportunity for the adaptive scheme to organically adjust to the function compared to the more rigid gridded nature of the FF. Additionally, much of the work to be carried out will require higher dimensional design spaces, and OLHS will likely scale better with dimensionality than FF. However, for continued testing and refinement in two dimensions, both initial sampling methods will continue to be used. This is because unlike OLHS, FF is repeatable and consistent on each subsequent run, thus eliminating a potential source of uncertainty.

B.2.4 Summary

In summary, for a given sample budget, the adaptive sampling method, as implemented in the code, achieves a better uncertainty measure than the single-stage space-filling methods in most cases, however, with an additional performance penalty for the generation time when compared to other methods. For analytical methods, this means it is often faster to run a single-stage sampling method with a higher sample budget than an adaptive scheme for a desired level of uncertainty. Nonetheless, if acquiring the sample takes significantly longer than it takes to identify new sample locations using the adaptive criteria, such as in the case of DSMC simulations, then adaptive methods are the faster approach for a given uncertainty measure. In general, this aligns with observations in work by other authors[86] which implies that the developed code is functioning as expected.
B.2. COMPARISON OF SAMPLING STRATEGIES

Figure B.4: RMSE versus the total run time of the sampling strategies

Figure B.5: RMSE versus the total run time of the sampling strategies assuming DSMC runtime of 1 hour per sample
B.3 Choosing a Support Radius

B.3.1 Introduction

A key element of the RBF method is the selection of the basis function. Many of these functions have additional parameters that allow them to be tuned to best suit the surrogate model. As described in Section 3.3.1.2, the basis function chosen for the work performed as part of this thesis was the Wendland’s $C^2$ function. This belongs to a family of functions that for a stated number of continuous derivatives $C^2_k$ in $n$ dimensions are compact and of minimal degree. These functions decay to zero at a given distance from the centre which is known as the support radius $R_s$, and is the tunable parameter for this family of functions. An example of a Wendland’s $C^2$ function for 2 dimensions is given in Equation 3.17.

Within the context of this surrogate model, the support radius controls the area of the domain over which a sample point has influence. If the radius is too small, then the interpolated surface may ‘sag’ in the regions between sample points. To avoid this, the support radius should be at least as large as the gaps between sample points. However, as the support radius gets large, a sample point may achieve global coverage of the domain and thus have an influence on all other sample points. If the support radius gets sufficiently large that all points have strong global influence then the system matrix $R$, which contains all pairwise combinations for the basis function (Equation 3.5), becomes poorly conditioned and thus accurate recovery of the model coefficients becomes difficult. The radius at which a sample obtains global coverage depends on its location, however, all samples will have global influence when the support radius is equal to the length of the longest diagonal of the normalized sample space hypercube, which for two dimensions/parameters is $1.41$.

B.3.2 Method

To determine how the support radius affected the generation and interpolation of the surrogate scenarios, a series of models were generated with normalized support radii of between 0.1 and 2 in steps of 0.1 using the four test cases described in Section 3.3.3.1 (Droplet Function, Franke Function, Aero-1 and Aero-2). For each test case, the four sampling strategies described in Section 3.3.2 were used (Fully Factorial (FF), Optimized Latin Hypercube Sampling (OLHS) and FF + Adaption and OLHS + Adaption), to identify if there was any dependency on the method when selecting the support radius. The generation process required a total of 100 samples each and the resulting models were compared against a 51x51 grid of each test case.

It should be noted that for a perfectly spaced 10x10 grid on a normalized sample space, a support radius of 0.1 is smaller than the spacing between sample sites and for some sampling strategies it may be much smaller than the maximum-minimum distance between observed. It has been included here to provide a worst-case comparison for the accuracy of the surrogate model. Similarly, even though a sample point may have global influence, its actual influence on another sample site may still be close to zero. This is certainly true for Wendland’s $C^2$ function if the distance between two points is only just smaller than the support radius. Thus while a normalized support radius of 2 is larger than the domain, it ensures that all samples have a non-zero influence over all other samples sites and thus provides a worst case situation for the conditioning of the system matrix $R$. 
B.3. CHOOSING A SUPPORT RADIUS

B.3.3 Results & Discussion

The RMSE error for each scenario is plotted against the support radii in Figure B.6, while Figure B.7 shows the maximum error for each case. As can be seen from the figures, for the majority of scenarios, both the RMSE and maximum errors appear to improve until the normalized support radius reaches about 1 (the length of the domain) where the uncertainty measures approach a lower limit. There is some variability between the different scenarios but in general, there is no substantial change or improvement in either the RMSE or Maximum Error for support radii above 1 for surrogate models generated with 100 samples. Thus to minimize the uncertainty in the model the support radius should be at least 1. The exceptions are the single-stage sampling strategies applied to the Droplet function and the Aero-1 test case. In the case of the former, and as was discussed in Appendix B.2, single-stage strategies are poorly suited to this type of function. In the case of the Aero-1 test case, 100 samples were not sufficient to accurately capture the steep and dramatic slope at the edge of the domain.

(a) Droplet Function

(b) Franke’s function

(c) AERO-1: Varying length and height

(d) AERO-2: Fixed dimensions and varying nose geometry

Figure B.6: Comparison of RMSE for each sampling strategies with varying support radii and a fixed sample budget of 100
A normalized support radius of 1, being the length of the domain, seems high and implies that any given sample point will likely influence a large proportion of the other samples in the domain. While this does not necessarily imply a strong influence over all these points, their influence is still non-zero which impacts the conditioning of the system matrix $R$. For a linear system, the condition number of the system matrix measures how sensitive the results are to perturbations in the input, such as errors in the sample data or roundoff errors made while calculating the solution. Plotting the reciprocal conditioning number, as is seen in Figure B.8, gives a sense of the smallest perturbation which will affect the most significant figure of the result. In this instance, a value of 1 indicates a well-conditioned matrix while a value of 0 indicates that the matrix is singular and not invertible. Thus a matrix that has a small reciprocal condition number is not well-conditioned.

The most prominent concern is round-off errors as a result of floating-point precision, given that the code is written in MATLAB with the matrix $R$ stored as a double-precision floating-point

![Figure B.7: Comparison of maximum error for each sampling strategies with varying support radii and a fixed sample budget of 100](image)
B.3. CHOOSING A SUPPORT RADIUS

Figure B.8: Reciprocal conditioning number for each sampling strategies with varying support radii and a fixed sample budget of 100

matrix. The relative floating-point accuracy (that is the smallest step size between numbers of the significand) is $2.2204 \times 10^{-16}$ ($2^{-15}$) which is significantly smaller than any of the reciprocal conditioning number observed in Figure B.8. This means that even though the system matrix is not well conditioned, the calculation of the coefficients is not unduly hindered by the floating-point accuracy even for a normalized support radius of 2.

For the work being done as part of this thesis, it was decided to use a support radius of 1.41 (the length of the diagonal) with the more general $R_s = \sqrt{n}$ for higher dimensions. As was seen from Figures B.6 & B.7, this ensures that the uncertainty values have attained their lower limits for the respective test cases and methods. Choosing 1.41, rather than 1 also provided a degree of consistency by enforcing global influence on all samples which in turn also ensured the good recovery of the surface even when there are few sample points, such as during the initial sampling stages of adaptive schemes.

B.3.4 Summary

In this section of work, the normalized support radius was varied from 0.1 to 2 to identify its effects on the uncertainty of the surrogate model. It was shown that the model uncertainty reached a lower limit when the support radius was above 1. It was also seen that while the systems matrices were in general poorly conditioned, for the scenarios and radii tested here the conditioning was more than sufficient to limit the impact of floating-point roundoff errors. A support radius of 1.41 (the length of the diagonal) was chosen for use in this project with the more general $R_s = \sqrt{n}$ for higher dimensions.
B.4 Order of the Polynomial

B.4.1 Introduction

In addition to the basis function, the radial basis function method has the option of including a polynomial as given in Equation 3.1. The rationale is that the polynomial can account for the trend of the data while the radial basis function provides a form of correction in areas that have a low agreement with that trend. Overall this should provide better agreement with the original function even in regions that have been poorly sampled.

For a polynomial with maximum degree \( d \) and number of variable \( N \) the number of monomial components of \( p(x) \), \( M \) is given by the binomial \( \binom{n+d}{n} \). Table B.2 shows the number of components of a polynomial with order 0 to 5 and the number of variables between 1 and 5. As can be seen from the table the number of components increases quickly with both the number of variables and the order of the polynomial. It is therefore important to know not only whether including a polynomial would be appropriate for the surrogate model in this setting but also what the order of polynomial should be used.

Table B.2: Number of Monomial components required for a polynomial of order 0 to 5 for a design space with 1 to 5 variables.

<table>
<thead>
<tr>
<th>Number of variables</th>
<th>Degree</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>3</td>
<td>62</td>
<td>10</td>
<td>15</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>4</td>
<td>10</td>
<td>20</td>
<td>35</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>5</td>
<td>15</td>
<td>35</td>
<td>70</td>
<td>126</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>6</td>
<td>21</td>
<td>56</td>
<td>126</td>
<td>252</td>
<td></td>
</tr>
</tbody>
</table>

B.4.2 Method

To examine whether a polynomial would be effective in this situation a set of scenarios were created using the four sampling strategies described in Section 3.3.2, (FF, OLHS and FF+ Adaption and OLHS+ Adaption) and the four test cases described in Section 3.3.3.1, (Droplet Function, Franke Function, Aero-1 and Aero-2). In each case, a series of surrogate models was generated, first without a polynomial and then using a polynomial with orders of 0 to 10. The generation process required a total of 100 samples each and used a support radius of \( R_S = 1.41 \). The resulting models were compared against a 51x51 grid of each test case.

B.4.3 Result & Discussion

The RMSE error for each scenario is plotted against the order of the polynomial in Figure B.9a while Figure B.9b show the maximum error for each case. As can be seen from Figures B.9a & B.9b there is no significant improvement in the RMSE or maximum error between including or excluding the polynomial for any of the scenarios tested. Some of the functions, notably Franke’s function, even experienced a worsening in the error for the higher-order polynomials. This could be as a result of over-fitting, particularly in the flatter regions of the domain.
An interesting side effect of introducing the higher-order polynomials is that under the adaptive scheme the criteria appears to prioritize the regions of high curvature first. This makes sense since polynomials of order 2 and above will contribute to the Laplacian of the model which plays a role in choosing the site of new samples. This would certainly be desirable in test cases such as Aero-1 which has a region of high curvature in one area of the domain. However, the improvements in fit around the slope appear to be entirely lost to the poor performance in the flat region of the domain. A test case like Aero-1 simply requires far more samples to achieve a consistent fit across the domain.

There is an increase in the run time for all scenario as the order of the polynomial is increased. For the single-stage methods, FF and OLHS, this was only marginal as the interpolation coefficients are only generated in a single execution once all samples have been collected. In the case of the adaptive methods, the coefficients must be regenerated each time a new set of samples is added which has a more significant impact on the runtime of the generator as the number of coefficients increases. This would be compounded further should the design space require more than two
parameters. As there is no improvement in the accuracy of the model, a higher-order polynomial would not be appropriate in this setting.

B.4.4 Summary

In this section it was demonstrated that for the test cases examined, there was no observed benefit to including an additional polynomial. Some methods suggest choosing the order of the polynomial using an optimization loop, but as has been demonstrated here, there is no significant change in performance with the change in the order of the polynomial or compared to not using one at all, thus an optimization loop would be superfluous. Including a polynomial is therefore not necessary for the method.
B.5 Sample Budget Easing to improve RMSE

B.5.1 Introduction

The sampling strategies described in Section 3.3.2 and tested in Appendix B.2 have thus far relied on a predefining sample budget before the model is generated. This works well for comparing methods on a known function, however, for an unknown function, it can be hard to predict what an appropriate sample budget might be and specifically, a sample budget that achieves the desired level of uncertainty in the surrogate model for the work being performed. An initial approach to selecting the sample budget could be to base the budget on previously generated surrogate models, however, as was seen in Figures B.1 & B.2 the surrogate models generated for each of the test cases had very different uncertainty measures for a given sample budget. By extension, if a certain level of uncertainty was required, the sample budget to achieve this varies significantly depending on the ‘black box’ function. While estimates could be made, from initial work with actual data it became apparent that a means of relaxing the budget was necessary. This section will address situations where easing the budget may be appropriate and explain how it was implemented in the code.

B.5.2 When to relax the budget and when to stop

As was discussed above it is sometimes necessary to relax the sample budget in order to achieve a better level of uncertainty in the surrogate model. However, for this to be effective it is necessary to identify when relaxing the budget would be appropriate. Figure B.11 shows the error histories of the four test cases described in Section 3.3.3.1, Droplet Function, Franke’s Function, Aero-1 and Aero-2, assuming that the surrogate model was intended to run for 100 samples, with around a third of the samples allocated to the initial sample as normal. The generator was then allowed to keep adding samples until a total of 500 sample points was achieved. This process was repeated with a range of initial total sample budgets, and it was generally observed that the error histories fell into three scenarios for a notional uncertainty requirement:

1. The uncertainty requirements were met with fewer sample points than the original budget.

2. The uncertainty requirements were met but with a greater number of sample points than the original budget.

3. The uncertainty requirements were never met.

In the case of scenario 1, the surrogate model should be terminated once the desired level of uncertainty is reached. It wastes computational time to continue and, in some cases, there is the risk of overfitting the surface, particularly if the initial estimate is far too high. If the uncertainty requirements have not yet been met, as is the case in scenario 2, then the budget should be relaxed (allowed to increase) to see if a better uncertainty can be obtained. However, care must also be taken to distinguish scenario 2 from scenario 3, which did not achieve the desired RMSE at all. Since scenario 3 could indicate a problem with the surrogate model, it is best to terminate the process at the original budget or as soon as possible.
The scenarios above thus provide two termination criteria for the generation process. A third
criterion is also included to prevent the system from running perpetually if neither of the previous
conditions are met, thus the termination criteria for easing the surrogate model generation are:

1. Terminate if the uncertainty requirements is met.
2. Terminate if the uncertainty requirements can never be met.
3. Terminate if maximum Budget met.

### B.5.3 Predicting the progress

Catching when the RMSE reaches a given error is a trivial process, but determining whether
a case is struggling to reach the desired level of uncertainty and should therefore terminate is
less straighforward. The simplest method to assess whether the RMSE is still approaching the
required level of uncertainty would be to look at the latest change of RMSE and maximum error.
However, as can be seen in Figure B.11, the error history can be erratic in places, particularly when
using the Leave One Out Cross-Point Validation (LOOCPV) uncertainty measure. A smoother
gradient can be obtained by using a set of the most recent RMSE values, however, even successful
schemes can have local regions with gradients near zero so there is no easy way to distinguish
between struggling schemes and those that should be left to run their course. This approach also
provides no guarantee that the desired uncertainty would be reached and is thus an unreliable
indication of the progress of the RMSE.

A broader view would be to look at the trend of the RMSE as a whole and extrapolate how many
samples would be needed to achieve the desired uncertainty. Based on this extrapolation, if no
real solution exists or the crossover point is an unachievably high number, the generation process
is terminated at the current Budget. The difficulty is that early in the generation process there

![Figure B.11: Comparison of error history for the four test cases over a sample range up to a sample budget of 500 and assuming an initial 'guess' of 100](image)
B.5. SAMPLE BUDGET EASING TO IMPROVE RMSE

Table B.3: $R^2$ Goodness-of-fit of a number curves to the RMSE values in $\log_{10}$ space for each test case

<table>
<thead>
<tr>
<th>Models</th>
<th>Droplet</th>
<th>Franke</th>
<th>Aero-1</th>
<th>Aero-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>linear $ax + b$</td>
<td>0.9040</td>
<td>0.9029</td>
<td>0.9772</td>
<td>0.5238</td>
</tr>
<tr>
<td>exponential $a \exp(bx)$</td>
<td>0.8421</td>
<td>0.8622</td>
<td>0.9675</td>
<td>0.5147</td>
</tr>
<tr>
<td>power law $ax^b$</td>
<td>0.9756</td>
<td>0.9894</td>
<td>0.9509</td>
<td>0.7784</td>
</tr>
<tr>
<td>power law + constant $ax^b + c$</td>
<td>0.9900</td>
<td>0.9955</td>
<td>0.9775</td>
<td>0.9748</td>
</tr>
</tbody>
</table>

may not be enough RMSE data points or the data is too noisy to produce a reliable extrapolation, especially if the uncertainty method to be used is LOOCPV. As stated above this method can be erratic and due to its high computational cost may only be run infrequently, such as only at the end of a batch of sample points (a batch being equal to the number of simulations that can be run in parallel on the system).

The first step, therefore, was to find a regression line that reliably describes the trend of the RMSE values for a range of test cases. The second step was to test the reliability of the extrapolation as sample points are added to the model and in particular when the number of RMSE values is low. As can be seen from Figure B.11 the trend of the RMSE of the four test cases is broadly exponential, though there were the occasional jumps, both up and down. Four curves were fitted to this RMSE history, these were a linear model ($ax + b$), an exponential model ($a \exp(bx)$), a power law ($ax^b$) and a power law with a constant offset ($ax^b + c$). These models were chosen because the trend of the RMSE history was observed to decay in a manner that was either linear or exponential, depending on the Test Case as can be seen in Figure B.11, Table B.3 shows the $R^2$ goodness-of-fit of each of the curves to the RMSE values for the four test cases and from the data it can be seen that the power law + constant ($ax^b + c$) consistently performed the best across all four test cases and is thus the prime candidate for the extrapolation curve.

B.5.4 Extrapolating the Uncertainty

The $R^2$ values in Table B.3 are based on fitting the curves to the complete RMSE history. In reality, these curves will need to be fitted to the reduced (but growing) data set to extrapolate the direction of the uncertainty of the surrogate model. However, as can be seen from Figure B.11 in some cases the trend of the data prior to the initial guess, in this case, 100 samples, can be a bit ‘noisy’, and may even suggest an increase in the RMSE before the trend settles into its downward course.

The ‘noisiness’ of the data early in the process may be due to a combination of factors. For instance, there are a limited number of evaluation sites available to the LOOCPV algorithm in the early part of the process, and thus as the number of sample points increases so too does the number of evaluation points and the downward trend becomes better established. Another consideration is that early in the process a sufficient number of samples may not have been reached to form an accurate representation of the surface, thus resulting in the lower level of precision and accuracy of the RMSE from the LOOCPV algorithm at low sample totals.
Figure B.12: Predicted sample budget to achieve an RMSE of at least $1 \times 10^{-3}$ and the associated $R^2$ values for these predictions for each of the test cases.
Since the ‘noisiness’ of the RMSE reduces as the sample size grows, it is useful to base the extrapolation process on the most recent samples as these will produce a more reliable prediction. The exact point at which this occurs is somewhat arbitrary but it was decided to discard the RMSE values for the first three-quarters of the original sample budget, in this case the RMSE values for all samples before the 75\(^{th}\) sample. Since this process is for assessing whether the generation process should continue adding sample points, the extrapolation only needs to be performed once the initial sample budget guess (in this case 100) has been surpassed.

Figure B.12 shows the evolution of the predicted sample budget required to achieve an uncertainty of less than \(1 \times 10^{-3}\) for the test cases (Droplet function B.12a, Franke’s function: B.12c, Aero-1: B.12e and Aero-2: B.12g) and the four regression models on the left with their associated \(R^2\) on the right (Droplet function B.12b, Franke’s function: B.12d, Aero-1: B.12f and Aero-2: B.12h). As can be seen from the figures, the power law with constant offset model consistently maintains the highest \(R^2\) value even with a small number of samples and is therefore achieving the best fit to the trends of the test case’s uncertainty. For a given test case, there can be quite a range of \(R^2\) values, especially when the number of samples is low. Low values of \(R^2\) indicate poor recovery of the trend of the error history and therefore the extrapolated values cannot be trusted. Accordingly, only extrapolations whose \(R^2\) values over 0.95 were considered when deciding whether to terminate or not.

The vertical and horizontal dashed lines in the predicted sample budget graphs in Figure B.12 identifies the size of the sample budget for each test case that was required to achieve an RMSE value of \(1 \times 10^{-3}\). Aero-1 is the exception as an RMSE value of less than \(1 \times 10^{-3}\) cannot be achieved with a sample budget less than 1500. In all other cases, though most easily demonstrated with the Droplet function (Figure B.12a), it can be seen that, despite the high \(R^2\) values, none of the models predict the correct total number initially. They do, however, appear to converge to the correct value as more samples are added. Importantly, the model does not need to perfectly predict the endpoint so long as it can show it that the uncertainty is achievable before the absolute maximum number of samples is reached. A good example of this is Aero-1, as it appears that the models converge towards a sample budget beyond the absolute maximum number of samples shown in Figure B.12e. In this case, it would make sense to stop the generation of the Aero-1 early as there is a clear trade off between generation time and the level of uncertainty in the model here.

The second termination condition would therefore be that when the power function with constant offset predicted that a sample budget greater than the absolute maximum was required to achieve the desired level of uncertainty in the model generation process should terminate.
APPENDIX B. NOTES ON THE DEVELOPMENT OF THE SURROGATE MODEL

B.5.5 The Break conditions

The error tripping process for the surrogate model therefore works as follows. A ‘guess’ is made of how many samples might be required to achieve the desired level of uncertainty in the model based on previously generated surrogate models with similar functions. This value is used to set the initial sample budget as well as the maximum sample budget. The maximum sample budget could be scaled depending on the confidence in the initial ‘guess’. For the purposes of testing here a value of 10 times the sample budget was used. The adaptive sampling would continue as normal but would break if one of the following conditions was met:

1. The RMSE drops below the desired uncertainty.
2. The predicted point at which the uncertainty is reached is greater than twice the maximum value.
3. The maximum number of samples is reached.

The RMSE was calculated using the LOOCPV algorithm as described in Section 3.3.4.2, since a set of gridded results for comparison will not be available during the actual live runs. Condition 2 was determined using a power function with a constant offset model and would only apply once the initial guess is surpassed and the model $R^2$ value was greater than 0.95.

This procedure was tested on all four test cases with the desired RMSE value of less than $1 \times 10^{-3}$ and using an initial guess of 100 samples. This implied an initial sample of 33 and an absolute maximum of 1000. Table B.4 shows the number of samples at the point the generation process was terminated, the uncertainty values for the initial 100 sample guess and the uncertainty values at the point the process stopped as well as the reason for stopping.

Table B.4: Comparison of the model uncertainty after an initial sample budget guess of 100 and after applying the easing process for a required RMSE value of less than $1 \times 10^{-3}$

<table>
<thead>
<tr>
<th>Test Case</th>
<th>No. Samples</th>
<th>Uncertainty after 100 Samples</th>
<th>Uncertainty at breakpoint</th>
<th>Reason for termination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Droplet</td>
<td>788</td>
<td>$8.89 \times 10^{-3}$</td>
<td>$4.89 \times 10^{-1}$</td>
<td>$8.98 \times 10^{-4}$</td>
</tr>
<tr>
<td>Franke</td>
<td>260</td>
<td>$5.01 \times 10^{-3}$</td>
<td>$2.20 \times 10^{-2}$</td>
<td>$9.69 \times 10^{-4}$</td>
</tr>
<tr>
<td>Aero-1</td>
<td>804</td>
<td>$1.03 \times 10^{-1}$</td>
<td>$7.59 \times 10^{-1}$</td>
<td>$1.72 \times 10^{-2}$</td>
</tr>
<tr>
<td>Aero-2</td>
<td>124</td>
<td>$1.52 \times 10^{-3}$</td>
<td>$7.27 \times 10^{-3}$</td>
<td>$9.36 \times 10^{-4}$</td>
</tr>
</tbody>
</table>

1 See list of break conditions

As can be seen from the table, by easing the restrictions on the budget, all but one case was able to achieve the desired uncertainty in the model. Aero-1 was the only test case not to achieve the required level of uncertainty. Figure B.12e demonstrates that the models had begun to converge on a value above 1500. The issue is whether the required uncertainty is worth the extensive simulation time. The testing performed here used panel methods to generate the data which runs incredibly quickly. Using DSMC simulation to gather the data would take significantly longer (estimated at least 7 days assuming 1 hour simulation time and 8 threads). Breaking the simulation at the point it did is therefore justifiable, as it has pushed the uncertainty down but recognized that the goal.
was not achievable given the constraints.

In all cases, the initial guess of 100 samples was not sufficient to obtain the desired uncertainty of less than $1 \times 10^{-3}$. A key issue here is that the more underestimated the guess the more rounds of adaptive sampling that need to be performed in order to reach the desired level of uncertainty which is less efficient. The Droplet function for instance required 688% more samples to achieve the required uncertainty all of which were added using adaptive sampling. The larger issue is that there is still no way of predicting how many samples a particular surface will need. For instance, the aerodynamic test cases, Aero-1 and Aero-2, required a significantly different number of samples and stopped generating for different reasons. This means that some form of budget easing is always going to be necessary.

### B.5.6 Summary

In this section, a method of dynamically easing the sample budget was presented incorporating a Power Law with a constant offset applied to the error history of the surrogate model in order to predict when the desired uncertainty threshold is reached.
C.1 The Angular Distribution of Gas Species

C.1.1 Introduction

This section will outline the derivation of the angular distribution of gases in a rarefied flow as used in Chapter 5.

C.1.2 Distribution of Velocities

The velocity of particles in a rarefied gas ($\mathbf{v}$) has two key components, the mean velocity of the bulk fluid ($\mathbf{v}_\infty$) and the internal velocity ($\mathbf{\tilde{v}}$) of the fluid, usually determined by its temperature. As the flow in VLEO is non-continuous (see Chapter 2) it is the random motion of the particles that brings them through the opening of a pit mounted on the side of the satellite. It is therefore useful to define a distribution that describes the angular probability of particles in rarefied gas relative to the velocity of the bulk fluid.

The direction of the bulk gas velocity ($\mathbf{v}_\infty$) is fixed by the macroscopic circumstances of the flow, in this instance a satellite passing through a region of VLEO. The magnitude and direction of the internal velocities $\mathbf{\tilde{v}}$ can be characterized by a thermal distribution function. Assuming that within the moving frame of the bulk fluid the particles have achieved kinematic equilibrium, then the Maxwell-Bolztmann distribution can be used to determine these velocities [45].

In three dimensions the Maxwell-Boltzmann distribution is given by Equation C.1. $u$, $v$ and $w$ in Equation C.1 are elements of the velocity vector ($\mathbf{v}$) and represent velocities in the $x$, $y$ and $z$ axis respectively. Here the Maxwell-Boltzmann distribution is expressed in terms of the most probable thermal speed $c'$ which is given by Equation C.2 where $m$ is the mass of the particle of the species, $T$ is the temperature of the gas and $k$ is the Boltzmann constant.

$$f(\mathbf{\tilde{v}})dudvdw = \left(\frac{1}{\pi c'^2}\right)^{\frac{3}{2}} \exp\left(-\frac{||\mathbf{\tilde{v}}||^2}{c'^2}\right)dudvdw$$  \hspace{1cm} (C.1)
\[ c' = \sqrt{\frac{2kT}{m}} \]  

Equation [C.1] can be rewritten in terms of the bulk velocity

\[ f(\mathbf{v})dudvdw = \left(\frac{1}{\pi c'^2}\right)^{\frac{3}{2}} \exp\left(-\frac{|\mathbf{v} - \mathbf{v}_\infty|^2}{c'^2}\right)dudvdw \]  

**C.1.3 Molecular Speed Ratio**

The molecular speed ratio can be defined as

\[ s = \frac{v}{c'} \]  

which also means the differential of \( v \) with respect to \( s \) is

\[ dv = dsc' \]  

So introducing the speed ratio into Equation [C.3]

\[ f(\mathbf{v})dudvdw = f(s)c'^3ds_udsv_ds_w \]  

\[ f(\mathbf{v})dudvdw = \left(\frac{1}{\pi}\right)^{\frac{3}{2}} \exp\left(-|s - |s_\infty||^2\right)ds_udsv_ds_w \]  

**C.1.4 Polar Form**

The next step is to convert the distribution in Equation [C.7] to a polar form with relation to direction of the bulk velocity in order to derive the angular distribution. Typically blockages in the flow will end on a straight edge with a finite length, such as the rear edges of the satellite or the rim of a recessed pit. Therefore it was decided to express the bulk velocity vector \( \mathbf{v} = [u, v, w] \) in terms of the cylindrical coordinates \([v_r, \theta, w]\) where \( v_r \) is the 2D velocity vector \([u, v]\) and \( w \) is parallel to the edge (see Figure [C.1]). \( \theta \) is the direction of \( v_r \) with respect to the bulk velocity \( v_\infty \).

\[ u = |v_r|\cos(\theta) \quad v = |v_r|\sin(\theta) \quad w = w \]
\[ s_u = |s_r|\cos(\theta) \quad s_v = |s_r|\sin(\theta) \quad s_w = s_w \]
C.1. THE ANGULAR DISTRIBUTION OF GAS SPECIES

Briefly assuming the $s_\infty$ is the molecular speed ratio of the bulk velocity $v_\infty$ and is positively defined in the $x$ direction.

$$|s - s_\infty|^2 = (s_u - |s_\infty|)^2 + s_v^2 + s_w^2 = s_u^2 - 2s_u |s_\infty| + |s_\infty|^2 + s_v^2 + s_w^2$$  \hspace{1cm} (C.8)

$$= |s_r|^2 \cos^2(\theta) - 2|s_\infty||s_r|\cos(\theta) + |s_\infty|^2 + |s_r|^2 \sin^2(\theta) + s_w^2$$  \hspace{1cm} (C.9)

$$= |s_r|^2 - 2|s_\infty||s_r|\cos(\theta) + |s_\infty|^2 + s_w^2$$  \hspace{1cm} (C.10)

Equation [C.7] therefore becomes

$$f(u)du dv dw = \left(\frac{1}{\pi}\right)^\frac{3}{2} \exp\left(2|s_r||s_\infty|\cos(\theta) - |s_r|^2 - |s_\infty|^2 - s_w^2\right)|s_r|ds_r|d\theta|dw$$  \hspace{1cm} (C.11)

C.1.5 Integrating the function

To derive the angular distribution $F(\theta)$, Equation [C.11] has to be integrated with respect to $|s_r|$ and $ds_w$.

$$F(\theta) = \int_{-\infty}^{\infty} \int_{0}^{\infty} \left(\frac{1}{\pi}\right)^\frac{3}{2} \exp\left(2|s_r||s_\infty|\cos(\theta) - |s_r|^2 - |s_\infty|^2 - s_w^2\right)|s_r|ds_r|d\theta|dw$$  \hspace{1cm} (C.12)

This can be split

$$F(\theta) = \left(\frac{1}{\pi}\right)^\frac{3}{2} \int_{0}^{\infty} |s_r| \exp\left(2|s_r||s_\infty|\cos(\theta) - |s_r|^2 - |s_\infty|^2\right|ds_r \int_{-\infty}^{\infty} \exp\left(-s_w^2\right)ds_w$$  \hspace{1cm} (C.13)

Integrating $s_w$

$$\int_{-\infty}^{\infty} \exp\left(-s_w^2\right)ds_w = \left[\frac{1}{2\sqrt{\pi}} \text{erf}(s_w)\right]_{-\infty}^{\infty}$$  \hspace{1cm} (C.14)
APPENDIX C. DERIVATIONS FOR THE DISTRIBUTIONS OF PARTICLES IN A GAS

Figure C.2: Angular probability distribution of the molecular flux relative to the bulk fluid velocity

Since $a > 0$, as $s_w \to \infty : \text{erf}(s_w) \to 1$ and $s_w \to -\infty : \text{erf}(s_w) \to -1$, therefore

$$\left[ \frac{1}{2} \sqrt{\pi} \text{erf}(s_w) \right]_{-\infty}^{\infty} = \frac{1}{2} \sqrt{\pi} (1 - (-1)) = \sqrt{\pi} \quad (C.15)$$

Returning this into Equation C.13

$$F(\theta) = \frac{1}{\pi} \int_{-\infty}^{\infty} |s_r| \exp \left( 2|s_r||s_\infty| \cos(\theta) - |s_r|^2 - |s_\infty|^2 \right) d|s_r| \quad (C.16)$$

Simplify the integral

$$F(\theta) = \exp \left( -|s_\infty|^2 \right) \frac{1}{\pi} \int_{-\infty}^{\infty} |s_r| \exp \left( 2|s_r||s_\infty| \cos(\theta) - |s_r|^2 \right) d|s_r| \quad (C.17)$$

\[= \frac{\exp \left( -|s_\infty|^2 \right)}{\pi} \left[ \frac{1}{2} \left( \sqrt{\pi}|s_\infty| \cos(\theta) \exp \left( |s_\infty|^2 \cos^2(\theta) \right) \text{erf}(|s_r| - |s_\infty| \cos(\theta)) \right. \right. \]
\[\left. \left. - \exp \left( 2|s_r||s_\infty| \cos(\theta) - |s_r|^2 \right) \right]_{0}^{\infty} \quad (C.18) \]

This therefore gives:

$$F(\theta) = \frac{1}{2\pi} \left( \sqrt{\pi}|s_\infty| \cos(\theta) \exp\left( -|s_\infty|^2 \sin^2(\theta) \right) \text{erf}(|s_\infty| \cos(\theta)) + 1 \right) + \exp(-|s_\infty|^2) \quad (C.19)$$

Equation C.20 represents the angular probability distribution of the molecular flux relative to the bulk fluid velocity ($v_\infty$). The distribution $F(\theta)$ is shown in Figure C.2 for a selection of $|s_\infty|$.
C.1.6 Angular Deviation

In the non-viscous environment of the upper atmosphere, the spread of the probability distributions of molecular flux in Figure C.2 determines how quickly the fluid will refill. As can be seen from this figure, the distribution appears to approach the form of a normal distribution. As $\theta \rightarrow 0$ Equation C.20 approaches:

$$F(\theta) = \frac{1}{2\pi} (\sqrt{\pi}|s_\infty| \exp(-|s_\infty|^2\theta^2)(\text{erf}|s_\infty| + 1) + \exp(-|s_\infty|^2))$$  \hspace{1cm} (C.21)

If $|s_\infty|$ is sufficiently large, $\text{erf}|s_\infty| \rightarrow 1$ and

$$F(\theta) = \frac{1}{2\pi} (\sqrt{\pi}|s_\infty| \exp(-|s_\infty|^2\theta^2)(1 + 1) + \exp(-|s_\infty|^2))$$  \hspace{1cm} (C.22)

$$F(\theta) = \frac{1}{\sqrt{\pi}}|s_\infty| \exp\left(-|s_\infty|^2\theta^2\right) + \frac{\exp(-|s_\infty|^2)}{2\sqrt{\pi}}$$  \hspace{1cm} (C.23)

Equation C.23 has the form of a normal distribution with an offset which tends to 0 as $|s_\infty|$ gets large. From this it can be inferred that the standard deviation for the angular distribution is

$$\sigma = \frac{1}{|s_\infty|\sqrt{2}}$$  \hspace{1cm} (C.24)

Since small angle theory was applied, the angular distribution can more accurately be defined as

$$\sigma_{AD} = \sin^{-1}\left(\frac{1}{|s_\infty|\sqrt{2}}\right)$$  \hspace{1cm} (C.25)

and can be understood to be the angular deviation of the particle flux from the bulk fluid velocity vector. Figure C.3 provides a comparison between $\sigma_{AD}$ and the numerically calculated standard deviation of $F(\theta)$ for a range of $|s_\infty|$. As can be seen from this figure, $F(\theta)$ converges towards a normal distribution as $|s_\infty|$ increases. When $|s_\infty| > 5$, the offset terms in Equation C.23 tends to zero, thus Equation C.23 also tends to a normal distribution.

The refill rate of rarefied gases was first described by Koppenwaller et al. in their work on concave corners, in which they introduced the concept of the molecular Mach angle ($\mu$) given by Equation C.26. Conceptually, the molecular Mach angle ($\mu$) is the angle a particle would take assuming its internal velocity ($\tilde{v}$) was equal to the most probable speed $c'$ in a direction perpendicular to the velocity vector of the bulk fluid. As with the angular standard deviation $\sigma_{AG}$, the molecular Mach angle represents the extent to which the particle flux deviates from the bulk velocity vector. A comparison of $\mu$ and $\sigma_{AD}$ are shown in Figure C.3, along with a result for the standard deviation calculated numerically from the angular distributions in Figure C.2.

$$\mu = \tan^{-1}\left(\frac{1}{|s_\infty|}\right)$$  \hspace{1cm} (C.26)
Figure C.3: Comparison of standard deviation of the angular probability distribution with $\sigma_{AD}$ and $\mu$. Numerical Standard deviation included for comparison.
C.2 Integrating the Wall distribution

C.2.0.1 Introduction

In the previous section an angular probability distribution for gas species in a rarefied flow was derived in Equation C.20. This was useful for understanding how the particles diverge from the bulk fluid velocity in a rarefied gas. This section will outline analytical solutions for the particle flux on the surface of panels within a pit mounted on the side of the satellite. These solutions will then be compared to a set of DSMC results to verify the equations.

C.2.0.2 Flux on a Wall

For a simple flat plate presented to a rarefied flow under steady state conditions, the particle flux for a given species is given by Equation C.27 where $N$ is the particle flux, $n_\infty$ is the number density of the species in the free-stream, $v_n$ is the velocity normal to the surface, $f(v)$ is the Maxwell-Boltzmann distribution for the velocity vector $v$ of the species and $X$ is a surface.

$$d^5N = n_\infty |v_n| f(v) du dv dw dX$$  \hspace{1cm} (C.27)

If the surface is fully exposed to the free-stream, this equation can be integrated over all inbound velocities normal to the surface and all velocities tangential to the surface to determine the fluency. However, in the case of a simple recessed pit, the bulk fluid particle can only reach the surfaces within the pit through the opening at the rim. This means that for particles to interact with a surface element $dX$, they must have an incoming radial velocity $v_r$ with an incident angle between $\eta_s$ and $\eta_r$, the angles from $dX$ to the Forward-facing rim (FFR) and Rearward-facing rim (RFR) of pit respectively (see Figure C.4). This means that the fluency will not be constant across surface $X$, so a flux distribution ($\Omega^N$) can be defined, given by the equation

$$\Omega^N_X = \frac{dN}{dX} = n_\infty \iiint_{\text{inbound}} f(v) |v_n| du dv dw$$  \hspace{1cm} (C.28)

Figure C.4: Diagram of a surface in a recessed pit
As in previous section, it is useful to express the velocity \( v = [u,v,w] \) in terms of cylindrical coordinates, this time \( [v_r, \eta, w] \). As before, \( v_r \) is the 2D radial velocity vector \([u,v] \) and \( w \) is parallel to the rim of the pit. \( \eta \) is the direction of vector \( v_r \) with respect to the normal of surface \( X \). Given this, the velocity normal to the surface \( v_n = v_r \cos(\eta) \). Equation [C.28] therefore becomes

\[
\Omega^N_X = \frac{dN}{dX} = n_\infty c' \int_{\rho_r}^{\eta_r} \int_{-\infty}^{\infty} f(v)|v_r|dw d|v_r|d\eta
\] (C.29)

\( f(v) \) was previously defined in Section C.1 with respect to the cylindrical coordinates \([v_r, \eta, w] \) by Equation [C.11] where \( \theta \) was the direction of \( v_r \) with respect to the bulk velocity \( v_\infty \). Inserting this into the Equation [C.29] and expressing the velocities in terms of the molecular speed ratio \( s \)

\[
\Omega^N_X = \frac{dN}{dX} = n_\infty \int_{\rho_r}^{\eta_r} \int_{-\infty}^{\infty} \left( \frac{1}{\pi} \right)^{\frac{3}{2}} \exp \left( 2s_r|s_\infty|\cos(\theta) - |s_r|^2 - |s_\infty|^2 - s_w^2 \right) |s_n||s_r|dw d|s_r|d\eta
\] (C.30)

Finally, integrating this over \( dw \) and \( d|s_r| \) produces

\[
\Omega^N_X = n_\infty \int_{\rho_r}^{\eta_r} \frac{c'}{2\pi} \left( \sqrt{\pi} |s_\infty|^2 \cos^2(\theta) + \frac{1}{2} \right) \exp \left( -|s_\infty|^2 \sin^2(\theta) \right) (\text{erf}(s_\infty \cos(\theta) + 1)
+ |s_\infty| \cos(\theta) \exp(-|s_\infty|^2)) \cos(\eta) d\eta
\] (C.31)

Equation [C.31] is the general form for the flux on a surface in the pit as a result of the gases entering the pit from the exterior. This equation does not account for reflections from other surfaces, though approximation can be made by applying this equation recursively with a reflection model.

### C.2.1 Element Analysis

It is not possible to integrate Equation [C.31] directly, so it is necessary to find an equivalent approximation. In order to find a suitable approximation, the first step was to analysis its elements in the form

\[
\Sigma^N_S = \frac{c'}{2\sqrt{\pi}} \int_{\theta_r}^{\pi/2} (F_1(\theta)F_2(\theta)F_3(\theta) + F_4(\theta)) \cos(\theta) d\theta
\] (C.32)

where

\[
F_1 = |s_\infty|^2 \cos^2(\theta) + \frac{1}{2}
\] (C.33)

\[
F_2 = \exp(-|s_\infty|^2 \sin^2(\theta))
\] (C.34)

\[
F_3 = \text{erf}(s_\infty \cos(\theta) + 1)
\] (C.35)

\[
F_4 = \frac{|s_\infty| \cos(\theta) \exp(-|s_\infty|^2)}{\sqrt{\pi}}
\] (C.36)

The elements \( F_1, F_2, F_3, \) & \( F_4 \) from Equations [C.33]-[C.36] are plotted between \(-180^\circ\) and \(180^\circ\) in Figures C.5a & C.5b for \( |s_\infty| = 1 \) and \( |s_\infty| = 5 \) respectively. As can be seen from Equation (C.35) element \( F_3 \) contains the error function \( \text{erf}(x) \) which allows element \( F_3(\theta) \) to act like a switch.
C.2. INTEGRATING THE WALL DISTRIBUTION

can be seen from the figures when $|s_\infty|$ is sufficiently large, $F_3$ is a positive constant around $\theta = 0$ but equal to zero around $\theta = \pm \pi$. $F_3(\theta)$ swaps between these two states at $\theta = \pm \pi/2$.

So as $\theta$ approaches $\pm \pi/2$, $F_1 F_2 F_3 \cos(\theta)$ approaches 0. Since $\theta_P \geq -\pi/2$, it can therefore be approximated that for $\pi/2 > \theta > -\pi/2$, $(\text{erf}(|s_\infty| \cos \theta) + 1) \approx (\text{erf}(|s_\infty|) + 1)$ or simply 2 when $|s_\infty|$ is sufficiently large. Equation C.48 can therefore be rewritten as

$$
\Omega_N^X = n_\infty \int_{0}^{\pi/2} \frac{c'}{2\pi} \left( \sqrt{\pi \left( |s_\infty|^2 \cos^2(\theta) + \frac{1}{2} \right)} \exp(-|s_\infty|^2 \sin^2(\theta))(\text{erf}(|s_\infty|) + 1) \\
+ |s_\infty| \cos(\theta) \exp(-|s_\infty|^2) \right) \cos(\theta) d\theta \quad (C.37)
$$

C.2.2 Verifying the Simplification

The next step is to check the accuracy of the simplification and determine when $|s_\infty|$ can be considered to be "sufficiently large" for the simplification to be valid. Figure C.6 shows the variation of the RMSE of the simplified equation, assuming $F_3 = 2$ or $F_3 = \text{erf}(|s_\infty|) + 1$, against the original distribution Equation C.37 for values of $|s_\infty|$ between 0 and 15. Figures C.6a & C.6b show the same data, just plotted on a linear graph and a logarithmic graph respectively for clarity.
APPENDIX C. DERIVATIONS FOR THE DISTRIBUTIONS OF PARTICLES IN A GAS

Figure C.6: RMSE of the simplified equation against the original distribution for values of $|s_\infty|$ between 0 and 15

The error function ($\text{erf}(x)$) approaches 1 as $x$ approaches $\infty$ (and -1 as $x$ approaches $-\infty$) but can be approximated as 1 for values of $x$ over 2. Similarly, as can be seen from Figure C.6, when $|s_\infty|$ is greater than 2, the RMSE for both approximations tends towards 0. However, by approximating $F_3$ to $\text{erf}(|s_\infty|) + 1$ rather than simply 2, ensures that the simplified equation maintains some accuracy even at low $|s_\infty|$.

So in general, the simplification of Equation C.31 to the form of Equation C.37 is valid for $|s_\infty| > 2$ but can still be applied to flows where $0 \leq |s_\infty| < 2$ with reduced accuracy. For rarefied flow in an orbital reference frame, it is reasonable to expect $|s_\infty|$ to be greater than 2 as is seen for most species in Table C.1. As can be seen by Table C.1, Hydrogen does have a $|s_\infty|$ less than 2, however, as can be also be seen from the table it makes up a very small percentage of the overall composition of the fluid and the results of the DSMC simulation to which the distribution is to be compared has low precision as can be seen by Figure 5.13h.

Table C.1: Mean speed ratio for the gas species and their associated angular deviations

<table>
<thead>
<tr>
<th>Gas Species</th>
<th>Composition [%]</th>
<th>$s_\infty$ [-]</th>
<th>$\sigma_{AD}$ [°]</th>
<th>$\mu$ [°]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atomic Oxygen ($O_1$)</td>
<td>56.20</td>
<td>7.41</td>
<td>5.48</td>
<td>7.68</td>
</tr>
<tr>
<td>Nitrogen ($N_2$)</td>
<td>39.89</td>
<td>9.80</td>
<td>4.14</td>
<td>5.82</td>
</tr>
<tr>
<td>Oxygen ($O_2$)</td>
<td>2.97</td>
<td>10.48</td>
<td>3.87</td>
<td>5.45</td>
</tr>
<tr>
<td>Helium ($He$)</td>
<td>0.59</td>
<td>3.71</td>
<td>11.00</td>
<td>15.10</td>
</tr>
<tr>
<td>Atomic Nitrogen ($N_1$)</td>
<td>0.28</td>
<td>6.93</td>
<td>5.85</td>
<td>8.21</td>
</tr>
<tr>
<td>Argon ($Ar$)</td>
<td>0.06</td>
<td>11.71</td>
<td>3.46</td>
<td>4.88</td>
</tr>
<tr>
<td>Hydrogen ($H$)</td>
<td>0.01</td>
<td>1.86</td>
<td>22.35</td>
<td>28.27</td>
</tr>
<tr>
<td>Mixture</td>
<td>100.00</td>
<td>8.53</td>
<td>4.76</td>
<td>6.69</td>
</tr>
</tbody>
</table>
C.2.3 Integrating the Simplification

Equation [C.37] is still a complex function to integrate, so it was necessary to break the equation down such that

\[ \Sigma^N_S = \frac{n_{\infty} c'}{2\sqrt{\pi}} [I + J + K] \]  
\[ (C.38) \]

where

\[ I = \int_{\theta_p}^{\pi/2} \sqrt{\pi} |s_{\infty}|^2 \cos^3(\theta) \exp(|s_{\infty}|^2 \cos^2(\theta)) (\text{erf}(|s_{\infty}|) + 1) d\theta \]  
\[ (C.39) \]

\[ J = \int_{\theta_p}^{\pi/2} \cos(\theta) \exp(|s_{\infty}|^2 \cos^2(\theta)) (\text{erf}(|s_{\infty}|) + 1) d\theta \]  
\[ (C.40) \]

\[ K = \int_{\theta_p}^{\pi/2} |s_{\infty}| \cos^2(\theta) d\theta \]  
\[ (C.41) \]

C.2.3.1 Integrating I

As element \( F_3 \) is a constant, this can now be factored out of the integral.

\[ I = \sqrt{\pi} (\text{erf}(|s_{\infty}|) + 1) \int_{\theta_p}^{\pi/2} |s_{\infty}|^2 \cos^3(\theta) \exp(|s_{\infty}|^2 \cos^2(\theta)) d\theta \]

\[ = \sqrt{\pi} (\text{erf}(|s_{\infty}|) + 1) \left[ \frac{\sqrt{\pi} (2|s_{\infty}|^2 - 1) \exp(|s_{\infty}|^2) (\text{erf}(|s_{\infty}|) \sin(\theta)) + 2|s_{\infty}| \sin(\theta) \exp(|s_{\infty}|^2 \cos^2(\theta))}{4|s_{\infty}|} \right]^{\pi/2}_{\theta_p} \]

\[ = \frac{\sqrt{\pi} (\text{erf}(|s_{\infty}|) + 1)}{4|s_{\infty}|} \left[ (\sqrt{\pi} (2|s_{\infty}|^2 - 1) \exp(|s_{\infty}|^2) (\text{erf}(|s_{\infty}|) - \text{erf}(|s_{\infty}| \sin(\theta_P))) + 2|s_{\infty}| (1 - \sin(\theta_P) \exp(|s_{\infty}|^2 \cos^2(\theta_P)))) \right] \]  
\[ = \sqrt{\pi} (\text{erf}(|s_{\infty}|) + 1) \left[ (\sqrt{\pi} (2|s_{\infty}|^2 - 1) \exp(|s_{\infty}|^2) (\text{erf}(|s_{\infty}|) - \text{erf}(|s_{\infty}| \sin(\theta_P))) + 2|s_{\infty}| (1 - \sin(\theta_P) \exp(|s_{\infty}|^2 \cos^2(\theta_P)))) \right] \]

C.2.3.2 Integrating J

As with element I, \( F_3 \) can be factored out of the integral.

\[ J = \frac{(\text{erf}(|s_{\infty}|) + 1) \sqrt{\pi}}{2} \int_{\theta_p}^{\pi/2} \cos(\theta) \exp(|s_{\infty}|^2 \cos^2(\theta)) d\theta \]  
\[ (C.42) \]

\[ J = \frac{(\text{erf}(|s_{\infty}|) + 1) \sqrt{\pi}}{2} \left[ \frac{\sqrt{\pi} \exp(|s_{\infty}|^2) \text{erf}(|s_{\infty}| \sin(\theta))}{2|s_{\infty}|} \right]^{\pi/2}_{\theta_p} \]  
\[ (C.43) \]

\[ J = \frac{(\text{erf}(|s_{\infty}|) + 1) \sqrt{\pi}}{2} \left( \frac{\sqrt{\pi} \exp(|s_{\infty}|^2) (\text{erf}(|s_{\infty}|) - \text{erf}(|s_{\infty}| \sin(\theta_P))))}{2|s_{\infty}|} \right) \]  
\[ (C.44) \]

\[ J = \frac{\text{erf}(|s_{\infty}|) + 1}{4|s_{\infty}|} \left( \pi \exp(|s_{\infty}|^2) (\text{erf}(|s_{\infty}|) - \text{erf}(|s_{\infty}| \sin(\theta_P)))) \right) \]  
\[ (C.45) \]
C.2.3.3 Integrating $K$

Unlike $I$ and $J$ integrating $K$ is relatively straight forward

$$K = \int_{\theta_p}^{\pi/2} |s_\infty| \cos^2(\theta) d\theta$$  \hspace{1cm} (C.46)

$$K = -\frac{1}{4} |s_\infty| (2\theta_p + \sin(2\theta_p) - \pi)$$  \hspace{1cm} (C.47)

C.2.3.4 Completing the Integration

Returning $I$, $J$ and $K$ into Equation C.38 and simplifying returns the final distribution

$$\Omega_X^N = \frac{n_\infty c'}{4\pi} \left( \text{erf}(|s_\infty|) + 1 \right) \left( |s_\infty| \text{erf}(|s_\infty|) - \text{erf}(|s_\infty| \sin(\theta)) \right)$$

$$-\sqrt{\pi} \sin(\theta) \exp(-|s_\infty|^2 \sin^2(\theta)) + \sqrt{\pi} \sin(\theta) \exp(-|s_\infty|^2)$$

$$-|s_\infty| \exp(-|s_\infty|^2/2) (2\theta + \sin(2\theta) - \pi)$$  \hspace{1cm} (C.48)

C.2.4 For large $|s_\infty|$

When $|s_\infty| > 5$ Equation C.48 can be simplified and rearranged to form Equation C.49. As can be seen from this new equation, the flux on the FFP of the pit can be expressed in terms of the maximum expected flux ($n_\infty u_\infty$) on a fully exposed surface. Furthermore, Equation C.49 also approximates the cumulative distribution function of a normal distribution. This is broadly to be expected since it was shown in Section C.1 that the angular distribution of particles can be approximated to a normal distribution with respect to $\sin(\theta)$. A cumulative distribution function is simply the probability density function integrated between $-\infty$ and $x$ which is similar to how Equation C.49 was obtained.

$$\Omega_X^N = n_\infty u_\infty \left( \frac{1}{2} \left( 1 - \text{erf}(|s_\infty| \sin(\theta)) \right) - \frac{\sin(\theta) \exp(-|s_\infty|^2 \sin^2(\theta))}{2 |s_\infty| \sqrt{\pi}} \right)$$  \hspace{1cm} (C.49)
REFERENCES


REFERENCES


[73] K. Chaloner et al. “Linked references are available on JSTOR for this article: Bayesian Experimental Design: A Review”. In: 10.3 (2016), pp. 273–304.


