

Comparing Active and Passive Just Noticeable Difference Thresholds for Stall Abruptness in Symmetric Stall

Master of Science Thesis

S. Bootsma

10 April 2024

Master of Science Thesis

Comparing Active and Passive Just Noticeable Difference Thresholds for Stall Abruptness in Symmetric Stall

Master of Science Thesis

MASTER OF SCIENCE THESIS

For the degree of Master of Science in Aerospace Engineering at Delft
University of Technology

S. Bootsma

10 April 2024

DELFT UNIVERSITY OF TECHNOLOGY
DEPARTMENT OF
CONTROL AND SIMULATION

The undersigned hereby certify that they have read and recommend to the Faculty of
Aerospace Engineering for acceptance a thesis entitled

COMPARING ACTIVE AND PASSIVE JUST
NOTICEABLE DIFFERENCE THRESHOLDS FOR
STALL ABRUPTNESS IN SYMMETRIC STALL

by

S. BOOTSMA

in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE AEROSPACE ENGINEERING

Dated: 10 April 2024

Supervisor(s):

dr.ir. D.M. Pool

dr.ir. C.C. de Visser

Committee Chair:

prof.dr.ir. M. Mulder

External Examiner:

dr.ir. M.F.M. Hoogreef

Preface

Dear reader,

Before you lies my masterpiece, or as most will call it, my master thesis. It is the crown jewel of my journey as a student at Delft University of Technology that will now be concluded after 7.5 years. My time as a student in Delft has been nothing short of a rollercoaster, a thrilling ride through time without knowing when the next looping will come, with the thesis being the final corkscrew before the brakes are applied to get back to reality. It is a journey that I will forever cherish, as it has allowed me to build and rebuild myself, my perseverance, and my resilience.

I could not have completed the journey without the help of numerous people along the way. I would like to start by expressing my gratitude to my thesis supervisors Daan and Coen. Especially Daan has been keeping a close eye on my research, helping me forward when needed, or simply asking the right question back when I arrived at his doorstep. Daan's care for his students is extraordinary, so I am grateful to have chosen him as a thesis supervisor. Furthermore, Coen's insight has proven to be extremely valuable, as he was able to see directly through the data what an underlying issue was. Working together with Daan and Coen in the stall task force has been a tremendous learning experience which was also very enjoyable. I would like to extend this gratitude to the other members of the stall task force as well, and wish the remaining members much wisdom in bringing this important subject further.

Furthermore, I would like to thank my parents and brother and sister for their support, not only during my thesis, but throughout my studies as a whole. You provided me with a place to escape my responsibilities for a weekend, such that I could reset and rest. Next to them, I would like to thank my friends who I've met during my student time. Your friendship has been incredible and I appreciate this deeply. A special thank you to the friends with whom I have studied and struggled together in room NB2.56. Thank you for the coffee breaks, discussions on our respective theses, Friday afternoon drinks, and unplanned other moments which might not have been the brightest ideas come the next morning. Sharing the thesis journey with you has been truly enjoyable.

Moreover, I would like to thank René and Olaf, as well as all the technicians who built and maintain the SIMONA Research Simulator for their help in setting up my experiment. Next

to them, I would like to thank the film score composers that wrote their amazing music, for providing me with a diverse music playlist that I could use throughout my entire thesis. Finally, I want to express my sincere gratitude to all participants who voluntarily signed up for my experiment. Their interest and great enthusiasm for my research has been a source of inspiration. It was a reminder of the importance of the work, making the upset recovery training even more safe than it already is.

To conclude this preface, I would like to reflect on where my student journey started. In my motivation letter to apply for the Control and Simulation master track, I wrote that I wanted to contribute towards a safer aviation industry. That is also what motivated me to start this thesis. The research that is presented in the following pages, for the first time, quantifies an accuracy that can be required to ensure that simulator-based stall training is truly effective. With the final words now being written in my thesis, I can proudly look back on my work with a firm belief that I did, indeed, contribute towards a safer aviation industry.

Sybren Bootsma
Delft, March 27, 2024

Table of Contents

| | |
|---|-------------|
| List of Figures | vii |
| List of Tables | xi |
| Nomenclature | xiii |
| 1 Introduction | 1 |
| I Scientific Paper | 3 |
| II Preliminary Thesis Report | 25 |
| 2 Background Information on Stall | 27 |
| 2-1 Aircraft Stall | 27 |
| 2-1-1 Aerodynamic Stall | 27 |
| 2-1-2 Stall Buffet | 29 |
| 2-2 Stall Models | 29 |
| 2-2-1 Stall Model Requirements | 29 |
| 2-2-2 Stall Modeling | 30 |
| 2-2-3 TU Delft Stall Model | 31 |
| 2-2-4 TU Delft stall buffet model | 35 |
| 2-3 Conclusions | 36 |
| 3 Human Differential Thresholds | 39 |
| 3-1 Just Noticeable Difference Thresholds | 39 |
| 3-2 Measurement strategies | 40 |
| 3-2-1 Method of constant stimuli | 40 |

| | | |
|----------|---|-----------|
| 3-2-2 | Staircase Procedure | 41 |
| 3-2-3 | Advanced Staircase Procedures | 42 |
| 3-2-4 | Parameter Estimation by Sequential Testing | 44 |
| 3-2-5 | Maximum-Likelihood Procedures | 45 |
| 3-3 | Biases | 45 |
| 3-4 | Passive vs Active Perception Thresholds | 46 |
| 3-5 | Conclusions | 47 |
| 4 | Stall in Flight Simulation Training Devices | 49 |
| 4-1 | Earlier Stall Simulator research | 49 |
| 4-1-1 | Work by Schroeder et al., (2014) | 49 |
| 4-1-2 | Work by Cunningham et al., (2019) | 50 |
| 4-1-3 | Work by Grant, Moszczynski, and Schroeder, (2018) | 51 |
| 4-1-4 | Other work | 51 |
| 4-2 | Earlier Stall Simulator Research at TU Delft | 51 |
| 4-2-1 | Work by Smets, De Visser, and Pool, (2019) | 52 |
| 4-2-2 | Work by Imbrechts, De Visser, and Pool, (2022) | 53 |
| 4-3 | Conclusions | 53 |
| 5 | Sensitivity Analysis | 55 |
| 5-1 | Method | 55 |
| 5-1-1 | Set up | 55 |
| 5-1-2 | Simulation | 57 |
| 5-2 | Results | 58 |
| 5-2-1 | Sensitivity of autopilot | 59 |
| 5-2-2 | Effects of different human control strategies | 60 |
| 5-3 | Conclusions | 61 |
| 6 | Experiment Setup | 63 |
| 6-1 | Experiment Hypotheses | 63 |
| 6-2 | Experiment Variables | 64 |
| 6-2-1 | Independent Variables | 64 |
| 6-2-2 | Dependent Variables | 64 |
| 6-2-3 | Control Variables | 65 |
| 6-3 | Other Experiment Design Variables | 66 |
| 6-3-1 | Participants | 67 |
| 6-3-2 | Condition Sequence | 67 |
| 6-3-3 | Experience Participants | 69 |
| 6-3-4 | Background Procedures during the Experiment | 69 |
| 6-4 | Final considerations | 70 |
| 6-4-1 | Two-alternative forced choice design | 71 |
| 6-4-2 | Staircase for passive experiment | 72 |
| 6-4-3 | Conditions active experiment | 72 |
| 6-4-4 | Flight Director | 73 |
| 6-5 | Conclusions | 73 |

| | |
|--|------------|
| III Appendices to Preliminary Thesis Report | 81 |
| A Additional Figures | 83 |
| A-1 Sensitivity analysis part 1 | 84 |
| A-1-1 States for different simulations centered around the stall | 84 |
| A-1-2 Delta of the states for different simulations centered around the stall with respect to the baseline | 89 |
| A-2 Sensitivity analysis part 2 | 94 |
| A-3 Experiment Design | 100 |
| B Additional Tables | 101 |
| B-1 Sensitivity Analysis | 101 |
| B-2 Sensitivity Analysis part 2 | 102 |
| B-2-1 Sensitivity with minimal gain $P_{\theta} = 0.2$ | 102 |
| B-2-2 Sensitivity with regular gain $P_{\theta} = 0.4$ | 103 |
| B-2-3 Sensitivity with maximum gain $P_{\theta} = 0.8$ | 106 |
| IV Appendices to Final Report | 109 |
| C Changes to Citation Stall Model | 111 |
| C-1 Separating the Model from Simulink simulation | 111 |
| C-2 Separating roll autopilot from stall autopilot | 111 |
| C-3 Moved calculations for stall autopilot to the controller block | 112 |
| C-4 Matching throttle of stall autopilot and trim values | 112 |
| C-5 Correcting undamped short period | 114 |
| C-6 Setting the Initial Conditions | 117 |
| C-7 Analysis of the Updated Model | 118 |
| D DUECA implementation of the experiment | 123 |
| D-1 The updated staircase procedure and ECI | 123 |
| D-2 How to use the ECI | 123 |
| D-3 SIMULINK Citation Model | 125 |
| E Experiment Document | 127 |
| F Individual results | 133 |

List of Figures

| | | |
|-----|--|----|
| 2-1 | Progress of trailing edge flow separation [12] | 27 |
| 2-2 | Three stall types [15] | 28 |
| 2-3 | Effect of a_1 on the lift curve and flow separation point (adapted from [26]) . . . | 32 |
| 2-4 | Effect of α^* on the lift curve and flow separation point (adapted from [26]) . . . | 33 |
| 2-5 | Effect of τ_1 on the lift curve and flow separation point (adapted from [26]) . . . | 33 |
| 2-6 | Effect of τ_2 on the lift curve and flow separation point (adapted from [26]) . . . | 34 |
| 2-7 | Schematic overview of the vertical stall buffet model as proposed in [24] | 36 |
| 3-1 | Example of a Psychometric Function [36] | 41 |
| 3-2 | Different results of a staircase procedure [38] | 42 |
| 3-3 | Example of a 2D/1U adaptive staircase procedure [36] | 43 |
| 3-4 | Weighted up/down staircase procedure with $\delta^-/\delta^+ = 1/3$ [39] | 44 |
| 3-5 | Comparison by Pentland [42] of several methods | 45 |
| 3-6 | Example of a Psychometric Function with different measurement strategies (2AFC, 3AFC, and 4AFC) [39] | 46 |
| 4-1 | Primary Flight Display used by Cunningham et al. [46] | 50 |
| 4-2 | Experiment procedures of [9] | 52 |
| 4-3 | Example of the results obtained by Smets, De Visser, and Pool [9] | 53 |
| 5-1 | The Cessna Citation II research aircraft PH-LAB [55] | 58 |
| 5-2 | Difference for all states between baseline and variations in $\alpha_{\text{threshold}}$ | 60 |

| | | |
|------|--|-----|
| 6-1 | SIMONA research simulator [9] | 66 |
| 6-2 | Experiment structure as designed by [9] | 70 |
| 6-3 | X and $\frac{dX}{dt}$ for the active experiment conditions as well as the maximum a_1 value for the passive experiment | 71 |
| | | |
| A-1 | Stall autopilot flow chart with steps during recovery phase [9] | 83 |
| A-2 | Influence of different settings of $\alpha_{\text{threshold}}$ | 84 |
| A-3 | Influence of different settings of θ_{ref} | 85 |
| A-4 | Influence of different settings of gain P_θ | 86 |
| A-5 | Influence of different settings of \dot{h}_e | 87 |
| A-6 | Influence of different settings of V_{TAS} | 88 |
| A-7 | Delta states of different settings of $\alpha_{\text{threshold}}$ with respect to the baseline | 89 |
| A-8 | Delta states of different settings of θ_{ref} with respect to the baseline | 90 |
| A-9 | Delta states of different settings of gain P_θ with respect to the baseline | 91 |
| A-10 | Delta states of different settings of \dot{h}_e with respect to the baseline | 92 |
| A-11 | Delta states of different settings of V_{TAS} with respect to the baseline | 93 |
| A-12 | Different a_1 settings with $\theta_{\text{ref}} = -3.5$ and gain $P_\theta = 0.2$ | 94 |
| A-13 | Different a_1 settings with $\theta_{\text{ref}} = -3.5$ and gain $P_\theta = 0.2$ | 95 |
| A-14 | Different a_1 settings with $\theta_{\text{ref}} = -3.5$ and gain $P_\theta = 0.2$ | 96 |
| A-15 | Different a_1 settings with $\theta_{\text{ref}} = 1.5$ and gain $P_\theta = 0.2$ | 97 |
| A-16 | Different a_1 settings with $\theta_{\text{ref}} = 1.5$ and gain $P_\theta = 0.4$ | 98 |
| A-17 | Different a_1 settings with $\theta_{\text{ref}} = 1.5$ and gain $P_\theta = 0.8$ | 99 |
| A-18 | Different a_1 settings of the active experiment, and the maximum a_1 setting from the passive experiment | 100 |
| | | |
| C-1 | The updated controller block with stall autopilot calculations included | 113 |
| C-2 | Stall autopilot throttle settings as by Smets, De Visser, and Pool [9] | 114 |
| C-3 | Updated auto-throttle of the stall autopilot | 115 |
| C-4 | Test run with no C_{m_q} with control input | 116 |
| C-5 | Test run with $C_{M_q} = -8.1826$ with control input | 117 |
| C-6 | Test run with $C_{M_q} = -20 \rightarrow -28$ with control input as shown in Figure C-5a | 118 |
| C-7 | Impulse response of the original $q - \delta_e$ system and the reduced system | 119 |
| C-8 | Simulation run with no inputs | 119 |
| C-9 | Simulation runs with all active values for a_1 with trimmed pitch rate q and elevator deflection δ_e | 120 |

| | | |
|------|---|-----|
| C-10 | Comparison of model by Smets et al [9] and the model used in this research for angle of attack and pitch angle | 120 |
| C-11 | Comparison of model by Smets et al [9] and the model used in this research for true airspeed and acceleration | 121 |
| C-12 | Comparison of model by Smets et al [9] and the model used in this research for pitch rate and altitude | 121 |
| C-13 | Comparison of model by Smets et al [9] and the model used in this research for flow separation point and elevator angle | 122 |
| D-1 | ECI for this experiment | 124 |
| D-2 | Flowdiagram of the DUECA modules in the Citation Stall folder | 126 |
| F-1 | Results of both the passive and active experiment for Subject 1, presented together with the percentages correct for the active experiment. | 133 |
| F-2 | Elevator inputs for all baseline runs from Subject 1, together with the pitch angle and flow separation point. | 134 |
| F-3 | Results of both the passive and active experiment for Subject 2, presented together with the percentages correct for the active experiment. | 134 |
| F-4 | Elevator inputs for all baseline runs from Subject 2, together with the pitch angle and flow separation point. | 134 |
| F-5 | Results of both the passive and active experiment for Subject 3, presented together with the percentages correct for the active experiment. | 135 |
| F-6 | Elevator inputs for all baseline runs from Subject 3, together with the pitch angle and flow separation point. | 135 |
| F-7 | Results of both the passive and active experiment for Subject 4, presented together with the percentages correct for the active experiment. | 136 |
| F-8 | Elevator inputs for all baseline runs from Subject 4, together with the pitch angle and flow separation point. | 136 |
| F-9 | Results of both the passive and active experiment for Subject 5, presented together with the percentages correct for the active experiment. | 137 |
| F-10 | Elevator inputs for all baseline runs from Subject 5, together with the pitch angle and flow separation point. | 137 |
| F-11 | Results of both the passive and active experiment for Subject 6, presented together with the percentages correct for the active experiment. | 138 |
| F-12 | Elevator inputs for all baseline runs from Subject 6, together with the pitch angle and flow separation point. | 138 |
| F-13 | Results of both the passive and active experiment for Subject 7, presented together with the percentages correct for the active experiment. | 139 |
| F-14 | Elevator inputs for all baseline runs from Subject 7, together with the pitch angle and flow separation point. | 139 |

| | | |
|------|--|-----|
| F-15 | Results of both the passive and active experiment for Subject 8, presented together with the percentages correct for the active experiment. | 140 |
| F-16 | Elevator inputs for all baseline runs from Subject 8, together with the pitch angle and flow separation point. | 140 |
| F-17 | Results of both the passive and active experiment for Subject 9, presented together with the percentages correct for the active experiment. | 141 |
| F-18 | Elevator inputs for all baseline runs from Subject 9, together with the pitch angle and flow separation point. | 141 |
| F-19 | Results of both the passive and active experiment for Subject 10, presented together with the percentages correct for the active experiment. | 142 |
| F-20 | Elevator inputs for all baseline runs from Subject 10, together with the pitch angle and flow separation point. | 142 |
| F-21 | Results of both the passive and active experiment for Subject 11, presented together with the percentages correct for the active experiment. | 143 |
| F-22 | Elevator inputs for all baseline runs from Subject 11, together with the pitch angle and flow separation point. | 143 |
| F-23 | Results of both the passive and active experiment for Subject 12, presented together with the percentages correct for the active experiment. | 144 |
| F-24 | Elevator inputs for all baseline runs from Subject 12, together with the pitch angle and flow separation point. | 144 |
| F-25 | Results of both the passive and active experiment for Subject 13, presented together with the percentages correct for the active experiment. | 145 |
| F-26 | Elevator inputs for all baseline runs from Subject 13, together with the pitch angle and flow separation point. | 145 |
| F-27 | Results of both the passive and active experiment for Subject 14, presented together with the percentages correct for the active experiment. | 146 |
| F-28 | Elevator inputs for all baseline runs from Subject 14, together with the pitch angle and flow separation point. | 146 |
| F-29 | Results of both the passive and active experiment for Subject 15, presented together with the percentages correct for the active experiment. | 147 |
| F-30 | Elevator inputs for all baseline runs from Subject 15, together with the pitch angle and flow separation point. | 147 |
| F-31 | Results of both the passive and active experiment for Subject 16, presented together with the percentages correct for the active experiment. | 148 |
| F-32 | Elevator inputs for all baseline runs from Subject 16, together with the pitch angle and flow separation point. | 148 |

List of Tables

| | | |
|------|--|----|
| 2-1 | Values used for creating Figures 2-3 to 2-6 | 34 |
| 2-2 | Proposed stall model parameters found by [25] | 35 |
| 2-3 | Stall buffet model parameters for the vertical direction | 36 |
| 2-4 | Stall buffet model parameters for the lateral direction | 36 |
| 4-1 | Abbreviated Stall Recovery Template [45] | 50 |
| 5-1 | Input parameters sensitivity analysis | 57 |
| 5-2 | Input parameters sensitivity analysis that would not benefit from percentage-wise step variation | 57 |
| 5-3 | Input parameters of sensitivity analysis part 2, with the Baseline values from [9] . | 57 |
| 5-4 | Different scenarios sensitivity analysis part 2 | 58 |
| 5-5 | Inputs for trim | 59 |
| 5-6 | Trim values for the simulation runs | 59 |
| 5-7 | VAF with respect to $\alpha_{\text{threshold}} = 0.28$ [rad] | 61 |
| 5-8 | VAF w.r.t. $\theta_{\text{ref_recovery}} = -0.5$ [deg] | 62 |
| 5-9 | VAF w.r.t. gain $P_{\theta} = 0.4$ | 62 |
| 5-10 | VAF w.r.t. $V_{\text{TAS}} = 86$ [m/s] | 62 |
| 6-1 | Active experiment settings for a_1 | 66 |
| 6-2 | Logic for experiment order for each participant, with N the number of other pilots | 68 |
| 6-3 | Latin Square design for active experiment conditions | 68 |
| 6-4 | Latin Square design with extra baseline conditions taken into account | 68 |

| | | |
|------|--|-----|
| B-1 | VAF with respect to $\dot{h}_e = -18 [m/s]$ | 101 |
| B-2 | Sensitivity of τ_1 with minimum θ_{ref} and minimum gain P_θ | 102 |
| B-3 | Sensitivity of a_1 with minimum θ_{ref} and minimum gain P_θ | 102 |
| B-4 | Sensitivity of τ_1 with maximum θ_{ref} and minimum gain P_θ | 103 |
| B-5 | Sensitivity of a_1 with maximum θ_{ref} and minimum gain P_θ | 103 |
| B-6 | Sensitivity of τ_1 with minimum θ_{ref} and normal gain P_θ | 104 |
| B-7 | Sensitivity of a_1 with minimum θ_{ref} and normal gain P_θ | 104 |
| B-8 | Sensitivity of τ_1 with maximum θ_{ref} and normal gain P_θ | 105 |
| B-9 | Sensitivity of a_1 with maximum θ_{ref} and normal gain P_θ | 105 |
| B-10 | Sensitivity of τ_1 with minimum θ_{ref} and maximum gain P_θ | 106 |
| B-11 | Sensitivity of a_1 with minimum θ_{ref} and maximum gain P_θ | 106 |
| B-12 | Sensitivity of τ_1 with maximum θ_{ref} and maximum gain P_θ | 107 |
| B-13 | Sensitivity of a_1 with maximum θ_{ref} and maximum gain P_θ | 107 |
| C-1 | Values for C'_{mq} as found by Van den Hoek, De Visser, and Pool [57] for different altitude and Mach number | 115 |
| C-2 | Additional trim values for q_0 and $\delta_{e_{extra}}$ for active experiment | 117 |

Nomenclature

Roman symbols

| | |
|-------|---|
| a_1 | A Kirchoff stall parameter related to the abruptness of the flow separation |
| b | Aircraft spanwidth |
| C | Coefficient |
| c | Constant |
| h | Height |
| I | Stimulus level |
| j | $\sqrt{-1}$ |
| N | Number of measurements |
| P | Gain |
| Q | Covariance matrix |
| q | Pitch rate |
| r | Yaw rate |
| T | Correct/Positive response |
| t | Time |
| V | Velocity |
| X | Position of flow separation point along chord ($X \in [0, 1]$) |
| x | Aircraft position along F_E x -axis |
| y | Aircraft position along F_E y -axis |
| y | Signal |
| z | Aircraft position along F_E z -axis |

Greek symbols

| | |
|------------|---|
| α | Angle of attack |
| α^* | A Kirchoff stall parameter representing the angle of attack for which $X_0 = 0.5$ |
| β | Angle of sideslip |
| δ | Control deflection |
| θ | Pitch angle |
| τ_1 | A Kirchoff stall parameter capturing the transient effects (time constant) |

| | |
|----------|--|
| τ_2 | A Kirchhoff stall parameter capturing the stall hysteresis (time constant) |
| ϕ | Roll angle |
| ψ | Correctness level |
| ψ | Yaw angle |
| ω | Angular frequency |

Subscripts

| | |
|------------|------------------------------------|
| 0 | Steady state conditions/trim point |
| <i>a</i> | Aileron |
| <i>D</i> | Drag coefficient |
| <i>e</i> | Elevator |
| <i>e</i> | Earth |
| <i>f</i> | Fuel |
| <i>L</i> | Lift coefficient |
| <i>l</i> | Roll moment coefficient |
| <i>m</i> | Pitch moment coefficient |
| <i>n</i> | Yaw moment coefficient |
| <i>r</i> | Rudder |
| <i>ref</i> | Reference |
| <i>s</i> | Stall |
| <i>T</i> | Thrust coefficient |
| <i>t</i> | Trim |
| <i>tas</i> | True airspeed |
| <i>y</i> | In Y_{F_b} direction |

Superscripts

| | |
|---|-----------------|
| · | Derivative |
| ^ | Estimate/mean |
| + | Upper Threshold |

Abbreviations

| | |
|--------|--|
| 1D/1U | 1 Down/1 Up Staircase procedure |
| 2AFC | 2-alternative forced choice |
| 2D/1U | 2 Down/1 Up Staircase procedure |
| DASMAT | Delft University Aircraft Simulation Model and Analysis Tool |
| EASA | European Union Aviation Safety Agency |
| FAA | Federal Aviation Administration |
| FSTD | Flight Simulation Training Device |
| ICAO | International Civil Aviation Organisation |
| JND | Just Noticeable Differences |
| LOC-I | Loss of Control In-Flight |
| ODE | Ordinary Differential Equation |
| PEST | Parameter Estimation by Sequential Testing |
| PFD | Primary Flight Display |
| SOC | Statement of Compliance |
| VAF | Variance Accounted For |

Chapter 1

Introduction

Loss of Control - Inflight (LOC-I) is currently the primary cause of fatal accidents in commercial aviation [1]. These accidents are often a result of pilots failing to prevent or recover from an upset. An upset is "*an undesired airplane state characterized by unintentional divergences from parameters normally experienced during operations. An airplane upset may involve pitch and/or bank angle divergences as well as inappropriate airspeeds for the conditions.*" [2]

One of the most predominant upsets is a stall. The importance of stall training was highlighted in 2009, when Colgan Air flight 3407, Turkish Airlines Flight 1951, and Air France flight 447 crashed. The National Transportation Safety Board wrote in their report the following probable cause of Colgan Air flight 3407: "The captain's inappropriate response to the activation of the stick shaker, which led to an aerodynamic stall from which the airplane did not recover." [3]. A similar cause can be found in the AF447 report: "Despite ... persistent symptoms, the crew never understood that they were stalling and consequently never applied a recovery manoeuvre." [4]. Finally, the Dutch Safety Board wrote in their report: "... the approach to stall recovery procedure was not executed properly, causing the aircraft to stall and crash." as one of the conclusions of the Turkish Airlines flight [5].

In response to these crashes, the International Civil Aviation Organization (ICAO) released a revision of the pilot training regulations which was implemented by the Federal Aviation Administration (FAA) and the European Aviation Safety Agency (EASA). Since 2019, they mandated pilots to follow a more extensive training program focusing on recovering from upsets. From 2019 onward, it is mandatory for commercial pilots to be trained on stall events different aircraft configurations, such as in take-off, clean, and landing configurations, as well as in different flight conditions [6]. These changes in training regulations required changes in the simulator part of the training as well, with increased requirements on the model fidelity beyond the critical angle of attack.

There is one thing that stands out on the updated simulator requirements: there is not a required accuracy or quantitative tolerance on the stall model fidelity given in the regulations. A flight simulation training device (FSTD) is certified for stalls by a subject matter expert who evaluates if the used model matches reality well enough [7], or in EASA's words: "for

each upset scenario, the recovery manoeuvre can be performed such that the FSTD does not exceed the FSTD training envelope, or when the envelope is exceeded, that the FSTD is within the realms of confidence in the simulation accuracy" [8]. This means that all around the world, stall training is done without truly knowing if the model is accurate enough.

One way of assessing how accurate the model needs to be, is to measure how sensitive a human is to detect differences. More specifically, how sensitive the human is to changes in the model. There has been earlier research [9, 10] on this topic, where so-called Just Noticeable Difference (JND) thresholds for variations in certain stall model parameters were determined. This research determined these thresholds while the participants were passive observers. The next step in this research is to determine if the thresholds are still applicable to a situation where the participant is actually flying themselves.

This master thesis report set up an experiment to determine to what extent JND thresholds found during a passive experiment are representative for JND thresholds found during an active pilot-in-command experiment. This work builds upon years of research on stall modeling and stall model verification work done by the Citation Stall Modeling Group of TU Delft's Faculty of Aerospace Engineering. Previous master theses focused on creating a working longitudinal stall model and thus far, several iterations of this model has been made. More recent work has focused on optimizing a lateral stall model as well [11]. Finally, the JND thresholds of several aspects of this model has been determined [9, 10]. This work continues the investigation.

This report is structured as follows. First, the scientific paper is shown in Part I. Then, the work leading up to the experiment is presented in the Preliminary Thesis Report in Part II, which was already graded for AE4020. In the Preliminary Thesis Report, background information on stall will be provided in Chapter 2, together with the introduction of the stall model used by the Citation Stall Modeling Group. After this, the Just Noticeable Difference threshold will be discussed as well as several measurement strategies to determine these thresholds. Furthermore, good practices of other stall simulation research will be shared in Chapter 4, in order to best prepare for this experiment. After this, a preliminary sensitivity analysis is discussed in order to get a good understanding of varying pilot behavior and its implications for this research. Finally, the experiment design is presented in Chapter 6. After the Preliminary Thesis Report, the relevant appendices for both the Preliminary Thesis Report and the scientific paper are given in Part III and IV, respectively.

Part I

Scientific Paper

Comparing Active and Passive Just Noticeable Difference Thresholds for Stall Abruptness in Symmetric Stall

Sybren Bootsma*

Faculty of Aerospace Engineering, Delft University of Technology, 2629HS Delft, the Netherlands

Aerodynamic stall has been a critical factor in recent aircraft crashes, leading to revisions in the regulations on the fidelity of stall models in flight simulation training devices. However, the updated regulations still lack a clearly defined accuracy required for effective pilot training. To determine the required accuracy, this research investigates how the Just Noticeable Difference (JND) thresholds for the stall abruptness parameter translate from a passive, observer task to an active flying task. An experiment was performed in the SIMONA Research Simulator with 16 active pilots, who performed two separate experiments. In one experiment, a stall autopilot flew the maneuver during which a staircase procedure was used to determine the passive JND threshold. The JND thresholds of $a_1^+ = 0.11 \pm 0.094$ found in this experiment were lower than the JND thresholds of $a_1^+ = 0.16 \pm 0.14$ found in a similar experiment by previous research. In the other experiment, the method of constant stimuli was used to determine the JND threshold for the active flying scenario. A psychometric curve, based on the Gaussian cumulative distribution function, was fitted using the combined responses of the participants. The resulting psychometric function of the active experiment lies entirely to the right of the passive psychometric function, and, when comparing the 75% thresholds for both experiments, the active threshold was found to be five times higher than the passive threshold. This indicates a decreased sensitivity to changes in stall abruptness when pilots are flying a stall themselves.

Nomenclature

Abbreviations

| | |
|------|--|
| 2AFC | two-alternative forced choice |
| 2D1U | two-down, 1-up |
| CDF | cumulative distribution function |
| FSTD | flight simulation training device |
| JND | just noticeable difference |
| RMS | root mean square |
| PEST | parameter estimation by sequential testing |

Roman Symbols

| | |
|-----------------|------------------------------|
| a_1 | stall abruptness |
| a_1^+ | upper threshold for a_1 |
| b | span |
| c | chord |
| C_D | drag coefficient |
| C_L | lift coefficient |
| C_l | roll moment coefficient |
| C_m | pitch moment coefficient |
| C_n | yaw moment coefficient |
| C_T | thrust coefficient |
| C_Y | side force coefficient |
| \dot{h}_e | Vertical speed |
| $K_{\{q,x,z\}}$ | pitch, heave, and surge gain |
| P | gain |

| | |
|-------|---|
| P | probability of correctness at stimuli level |
| p | statistical significance |
| p | roll rate |
| q | pitch rate |
| r | yaw rate |
| V | velocity |
| X | flow separation point |
| X_0 | steady flow separation point |

Greek Symbols

| | |
|--------------------------|---|
| α | angle of attack |
| α^* | angle of attack for which $X_0 = 0.5$ |
| β | angle of side slip |
| δ_a | aileron deflection |
| δ_e | elevator deflection |
| δ_r | rudder deflection |
| $\zeta_{\{q,x,z\}}$ | pitch, heave, and surge damping coefficient |
| θ | pitch angle |
| μ | mean |
| σ | variance |
| τ_1 | stall time delay constant |
| τ_2 | stall hysteresis time constant |
| φ | stimuli level |
| $\omega_{b_{\{q,x,z\}}}$ | pitch, heave, and surge break frequency |
| $\omega_{n_{\{q,x,z\}}}$ | pitch, heave, and surge natural frequency |

*MSc Student, Department of Control & Simulation,

I. Introduction

Loss of Control - Inflight (LOC-I) is currently the primary cause of fatal accidents in commercial aviation [1]. These accidents are often a result of pilots failing to prevent, or recover from, an upset. One of these upsets is the stall, a situation where the critical angle of attack is exceeded, which leads to a sudden loss of lift [2]. Aerodynamic stall has been a primary factor in several recent aircraft crashes [3–5]. Since then, the International Civil Aviation Organization has updated the regulations regarding upset recoveries, which has come into effect in 2019 [2, 6, 7]. Numerous aspects of the upset prevention and recovery training were updated, including the required fidelity of the stall model used in flight simulation training devices (FSTDs). An FSTD is currently certified for stalls by a subject matter expert who evaluates if the used model matches reality well enough [8] or as mentioned in EASA's regulations: "*for each upset scenario, the recovery manoeuvre can be performed such that the FSTD does not exceed the FSTD training envelope, or when the envelope is exceeded, that the FSTD is within the realms of confidence in the simulation accuracy*" [9].

Previous efforts have validated that stall models match reality sufficiently through pilot-in-the-loop simulations. The work by Schroeder et al. [10] evaluated several stall models in a Level D B737-800 simulator. The four stall models used during the experiment were tested by several pilots experienced in flying stalls on the B737. The models were an old model from before the 2019 regulations update, an updated model by Boeing for the 2019 regulations, and two models based on scaled wind tunnels tests, computational aerodynamics, and expert opinions. When asked if the stall models could be used for training, the participants "*somewhat agreed*" [10]. Furthermore, they found no significant difference in recovery performance between the models. Finally, they concluded that effective stall models could be developed based on wind tunnel data, computational aerodynamics, and expert pilot opinions.

Other work by Grant et al. [11] used three different models for their comparison. They used a nominal model and compared it with an extreme and a mild model, both of which were derived from the nominal model. The extreme and mild model were formed by changing the aerodynamic parameters. Furthermore, Grant et al. [11] increased the roll-off and stall buffet for the extreme model, and decreased their intensities for the mild model. They found no significant difference in recovery performance of participants between the different models, similarly to Schroeder et al. [10]. Moreover, their participants indicated that each of the models developed were acceptable for stall training.

A first step in looking at the sensitivity of pilots was taken by Cunningham et al. [12], who evaluated the sensitivity to changes with regards to several factors of their stall model. These factors included, among others, stall asymmetries, control effectiveness, and dynamic stability. They asked participants to rate the significance of the changes with respect to these factors on a scale of 0-9, where 0 meant no differences and 9 large differences. Their participants indicated that the stick pusher and stall asymmetries have the most significant changes in the simulation, the dynamic stability as function of angle of attack rate the least.

A similar effort was performed by Smets et al. [13] and Imbrechts et al. [14], who investigated the Just Noticeable Difference (JND) thresholds of key parameters of the stall model and the stall buffet model, respectively. This was done through simulator experiments using a staircase procedure where the baseline model was compared to a model that contained the updated parameter. However, participants were not in control of the approach-to-stall and stall recovery, since a specifically designed stall autopilot flew the procedure. This means that the participants could use their full attention for detecting differences between the baseline and adjusted model.

This leaves the question what happens to pilots' JND thresholds when they fly the aircraft themselves. For this, other research at TU Delft [15, 16] provides insights. Hosman and van der Vaart [15] investigated how passive vestibular motion perception thresholds change when additional mental workload is added. They found an increase in threshold of 26% to 266%, depending on the kind of mental workload added. Valente Pais et al. [16] looked into the absolute threshold in pitch motion during an active control task and a passive observing task and found that the absolute threshold was 60% higher in the active control experiment. Both works indicate that a higher threshold is found when additional mental workload or control inputs are required.

Building upon these findings, this research investigates how pilots' JND thresholds differ between an observing task and an active control task, for the stall abruptness parameter of a Kirchhoff stall model. This question is answered by performing a pilot-in-the-loop experiment, where the upper threshold for both a passive autopilot controlled scenario and an active pilot-in-command scenario are determined. These scenarios will be referred to as the passive and active experiment, respectively. Both experiments were conducted in the SIMONA Research Simulator at the Faculty of Aerospace Engineering at TU Delft. In the experiment, 16 pilots experienced a symmetric stall maneuver at around 18,000 feet. A Parameter Estimation by Sequential Testing (PEST) staircase procedure [17] was used for the autopilot scenario and the method of constant stimuli [18] was used for the pilot-in-command scenario. During each test run, participants compared the baseline stall model to a model with an increased abruptness in the flow separation and

were asked to indicate which model showed a more abrupt response. Based on the answers given, the threshold and psychometric function were determined for both experiments, which gives insight in how the JND threshold shifts from the passive observing task to the active control task.

This paper is structured as follows. First, Section II provides additional information on the aerodynamic stall model used in this research, as well as the changes made to the model to allow active pilot control. Section III provides the hypotheses for this research and describes the experiment set up used to determine the active and passive thresholds for changes in stall abruptness parameter of the Kirchhoff stall model, followed by the results of the experiment in Section IV. Finally, the results are discussed in Section V, which is followed by the conclusions in Section VI.

II. Aerodynamic Stall Model

A. Stall Model based on Kirchhoff's Theory of Flow Separation

The aerodynamic stall model used in this experiment has been developed by earlier research at TU Delft [19, 20] and is based on Kirchhoff's Theory of Flow Separation, as first described by Fischenberg [21]. This model is centered around the calculation of a flow separation point, X . The flow separation point is defined to be 1 when the flow is fully attached to the wing and 0 when the flow is fully separated. X is dependent on many factors, such as the hysteresis factor, which captures the circulation and boundary layer effects through Kirchhoff's theory. By combining this with the Wagner or Theodorsen function to capture the unsteady aerodynamics, Fischenberg [21] made the flow separation point time-dependent. Finally, X can be estimated from the steady flow separation point, when combined with the previous aspects. This ultimately results in the nonlinear ordinary differential equation that can be used to dynamically model the flow separation, as can be seen in Equation (1). For more information regarding the model, see Van Ingen et al. [20] or Fischenberg [21].

$$\tau_1 \frac{dX}{dt} + X = \frac{1}{2} \{1 - \tanh(a_1(\alpha - \tau_2 \dot{\alpha} - \alpha^*))\} \quad (1)$$

In Equation (1), the different parameters all capture different characteristics of the stall. a_1 [-] is indicative of the stall abruptness, where a higher value means a more abrupt drop in lift and a sudden and quick flow separation. This is the parameter that will be varied to measure the threshold for changes in stall abruptness. Furthermore, α^* [rad] is the angle of attack for which the flow separation point $X = 0.5$. τ_1 [s] and τ_2 [s] are both time constants that capture the transient and hysteresis effects of the stall, respectively [21–24]. The flow separation point X is used in the longitudinal aircraft model of the Cessna Citation II, the research aircraft of the Faculty of Aerospace Engineering at TU Delft*. This model is obtained from Van Ingen et al. [20], with the longitudinal model as given by Equations (2) - (4), with the exception of the C_{m_q} term indicated in red which will be discussed in Section II.B. The coefficients are given in Table 1. The lateral model used in this research is given by Van Ingen et al. [20] as well. Furthermore, the adaptive, X -dependent stall buffet model as described by Van Horssen et al. [19] is used.

$$C_L = C_{L_0} + C_{L_\alpha} \left(\frac{1 + \sqrt{X}}{2} \right)^2 \alpha + C_{L_{\alpha^2}} (\alpha - 6)_+^2 \quad (2)$$

$$C_D = C_{D_0} + C_{D_\alpha} \alpha + C_{D_{\delta_e}} \delta_e + C_{D_X} (1 - X) + C_{D_{C_T}} C_T \quad (3)$$

$$C_m = C_{m_0} + C_{m_\alpha} \alpha + C_{m_q} \frac{qc}{V} + C_{m_{X\delta_e}} \max\left(\frac{1}{2}, X\right) \delta_e + C_{m_{C_T}} C_T \quad (4)$$

B. Additional Pitch Damping

During the initial testing phase of the experiment, it was found that, when manually flying the stall model, the pitch rate q was not properly damped, despite the Cessna Citation II having a strongly damped short period [25]. Consequently, the eigenvalues of the $q - \delta_e$ model in the stall regime were determined by reducing the states and finding the eigenvalues in MATLAB. From this, a positive pair of complex eigenvalues were found, with a value of $2.42 \cdot 10^{-3} \pm 0.192i$. The Cessna Citation II should have a damped short period with eigenvalues of $-3.9161 \cdot 10^{-2} \pm 3.7971 \cdot 10^{-2}i$ in nominal flight conditions [25].

*<https://cs.lr.tudelft.nl/citation/>

Table 1 Proposed stall model parameters from [20].

| Name | Value | Unit | Name | Value | Unit | Name | Value | Unit |
|--------------------|---------|-------|--------------------|---------|------|----------------------|---------|------|
| a_1 | 27.6711 | [-] | $C_{D(1-x)}$ | 0.0732 | [-] | C_{l_r} | 0.1412 | [-] |
| α^* | 0.2084 | [rad] | $C_{D_{C_T}}$ | 0.3788 | [-] | $C_{l_{\delta a}}$ | -0.0853 | [-] |
| τ_1 | 0.2547 | [s] | C_{Y_0} | 0.0032 | [-] | C_{m_0} | 0.0183 | [-] |
| τ_2 | 0.0176 | [s] | C_{Y_β} | -0.5222 | [-] | C_{m_α} | -0.5683 | [-] |
| C_{L_0} | 0.1758 | [-] | C_{Y_p} | -0.5000 | [-] | $C_{m_{\delta e x}}$ | -1.0230 | [-] |
| C_{L_α} | 4.6605 | [-] | C_{Y_r} | 0.8971 | [-] | $C_{m_{C_T}}$ | 0.1443 | [-] |
| $C_{L_{\alpha^2}}$ | 10.7753 | [-] | $C_{Y_{\delta a}}$ | -0.2932 | [-] | C_{n_0} | 0.0013 | [-] |
| C_{D_0} | 0.0046 | [-] | C_{l_0} | -0.0017 | [-] | C_{n_β} | 0.0804 | [-] |
| C_{D_α} | 0.2372 | [-] | C_{l_β} | -0.0454 | [-] | C_{n_r} | -0.0496 | [-] |
| $C_{D_{\delta e}}$ | -0.1857 | [-] | C_{l_p} | -0.1340 | [-] | $C_{n_{\delta r}}$ | 0.0492 | [-] |

Consequently, a pitch damping coefficient C_{m_q} was added to the pitch moment equation, since the model by Van Ingen et al. [20] did not contain this term, resulting in the equation as seen in Equation (4) with additional C_{m_q} shown in red. Initially, a C_{m_q} value was estimated based on the work by Van den Hoek et al. [26], who identified a nominal flight envelope model. Their work allowed to interpolate a C_{m_q} value based on altitude and Mach number, resulting in $C_{m_q} = -8.1826$. This updated model was verified through pilot-in-the-loop evaluations with an experienced Cessna Citation II test pilot. These experiments verified that the damping that was included, was still not sufficient. Although the addition of C_{m_q} damped the oscillatory pitch rate, the damping was not as strong as the behavior found in the Cessna Citation II, according to the test pilot. Consequently, the value of C_{m_q} was further increased based on offline simulations, which can be seen in Figure 1. Here, the control inputs of the pilot-in-the-loop test (with $C_{m_q} = -8.1826$) were used to analyze the effects of an increase in C_{m_q} . Figure 1 shows that $C_{m_q} = -22$ sufficiently damps the found behavior. The updated model was again verified with pilot-in-the-loop evaluations, which confirmed that the updated model represented the pitch behavior of the Cessna Citation II in the approach to stall and post-stall flight.

The short period eigenvalues of the updated model in the stall regime were found to be $-7.19 \cdot 10^{-3} \pm 0.142i$. Despite the fact that these values are half an order of magnitude away from the values given by Mulder et al. [25] in nominal flight conditions, the confirmation through the pilot-in-the-loop simulations is deemed acceptable for this research. Further integration of the C_{m_q} within the stall model is outside the scope of this research and left as a crucial recommendation for future research. This analysis is described in more detail in Appendix C.

As a final verification step, a comparison of an autopilot test run between these updated model and the model used by Smets et al. [13] was performed. This was done to ensure that the differences between the models did not result in different threshold values. The results for the comparison between the pitch angle for the two models can be seen in Figure 4. Here, it can be seen that the difference in pitch between the baseline and the JND threshold found by Smets et al. [13] is similar to the difference between the baseline and the JND threshold setting for a_1 with the updated model. This was found to be true for other key states as well, which can be seen in Appendix C. Therefore, it is assumed that the thresholds found in this research can be compared to the passive threshold found by Smets et al. [13].

C. Stall autopilot sensitivity analysis

To be able to compare the upper thresholds of the stall abruptness for both the active and passive scenario, both scenarios must be flown as similar as possible. On the other hand, letting participants focus too much on following an exact flight path can lead to distractions from the actual experiment. Consequently, a balance needed to be found between flying both scenarios similarly versus giving participants the freedom to fly as they are trained.

To determine which part of the maneuver needed to be restricted and could be left up to the participants, an offline sensitivity analysis was performed. The stall autopilot decision tree as used by Smets et al. [13] was taken as a reference and a sensitivity analysis was performed on the key parameters of the autopilot. This mimicked the differences expected in the control behavior of real participants. The key parameters of the stall autopilot that were varied are:

- The threshold $\alpha_{\text{threshold}}$ at which the recovery procedure is initiated, currently set at $\alpha_{\text{threshold}} = 16.04^\circ$
- The reference angle θ_{ref} during the stall recovery phase, currently set at $\theta_{\text{ref}} = -0.5^\circ$
- The controller gain P_θ of the reference angle θ_{ref} , currently set at $P_\theta = 0.4$ [-]

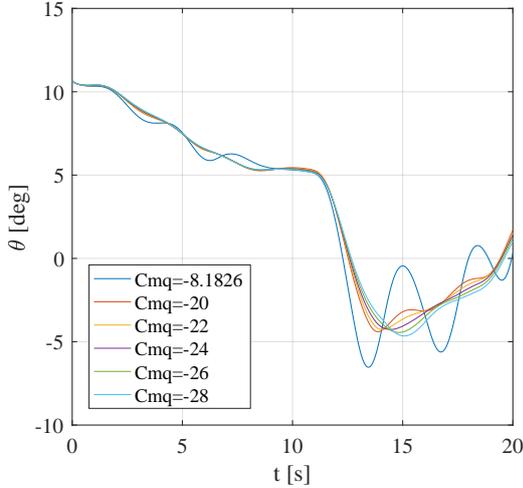


Fig. 1 Pitch angle data from the pilot-in-the-loop experiment in SIMONA with $C_{mq} = -8.1826$ and several offline simulations for different C_{mq} using the control inputs from the experiment.

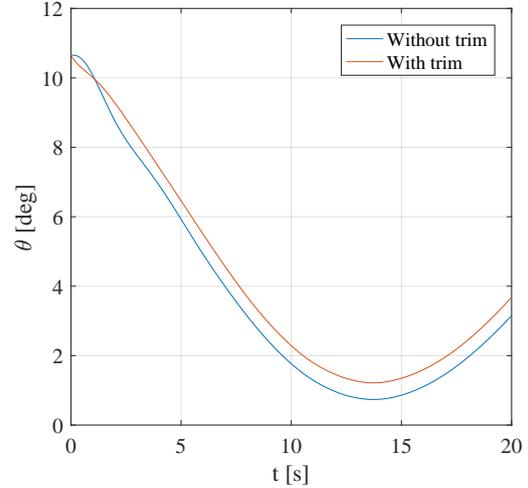


Fig. 2 Pitch angle for trimmed and untrimmed model, including the updated model with $C_{mq} = -22$.

- The threshold on \dot{h}_e when full thrust is applied, currently set at $\dot{h}_e = -18$ [m/s]
- The threshold on V_{tas} when the reference angle switches back to 10° to go back to the original flight path, currently set at $V_{tas} = 86$ [m/s]

The threshold for starting the recovery procedure $\alpha_{\text{threshold}}$ was used as a starting point for the variation. The starting angle of attack was varied initially, to see what a valid range for the sensitivity could be. Variations larger than 18.5° were deemed unrealistic, since there is already full flow separation at 16° resulting in an unrealistic recovery scenario. Furthermore, a $\alpha_{\text{threshold}}$ smaller than 13.5 degrees would not lead to a fully developed stall. Consequently, the variations for the sensitivity analysis were set to vary the parameters with $\pm 15\%$, resulting in a range of $13.6^\circ - 18.4^\circ$ for $\alpha_{\text{threshold}}$. As a result, the other parameters were also varied with $\pm 15\%$ for the analysis to allow for a fair comparison between the parameters. The variations in important aircraft states such as pitch, pitch rate, velocity, flow separation point, elevator input, and angle of attack were analyzed. These states were compared to the variations in the states that were found with the upper and lower a_1 thresholds of Smets et al. [13].

It was found that the only parameter that would lead to noticeable variations in the states, is the threshold for starting the recovery procedure, $\alpha_{\text{threshold}}$ [$^\circ$], which can be seen in Figure 3. This figure shows that the difference in velocity, pitch angle, and elevator input lie beyond the upper and lower a_1 thresholds found by Smets et al. [13] and are therefore likely noticeable for participants. The variations of the other parameters resulted in differences in the states that fell within the upper and lower a_1 thresholds. Hence, it is concluded that allowing participants to fly freely the recovery process, while restricting the stall entry conditions is the optimal balance. This approach limits the constraints put on participants while ensuring they can notice differences between the stall abruptness.

During testing of the experiment for the active scenario, it was found that the scenario was not properly trimmed, resulting in an oscillatory pitch behavior. Consequently, each active scenario was trimmed manually, and the initial conditions for pitch rate q were set as well. This proved effectively in reducing the initial oscillatory pitch behavior, as can be seen in Figure 2. It was chosen to omit this additional trim for the passive experiment, as the stall autopilot is able to keep the pitch attitude at the desired angle without oscillations.

III. Method

A. Hypotheses

Currently, it is unknown how the JND thresholds for the parameters of the Kirchhoff stall model translate from a passive, observing task to a pilot-in-command task. Hence, the research question for this research is:

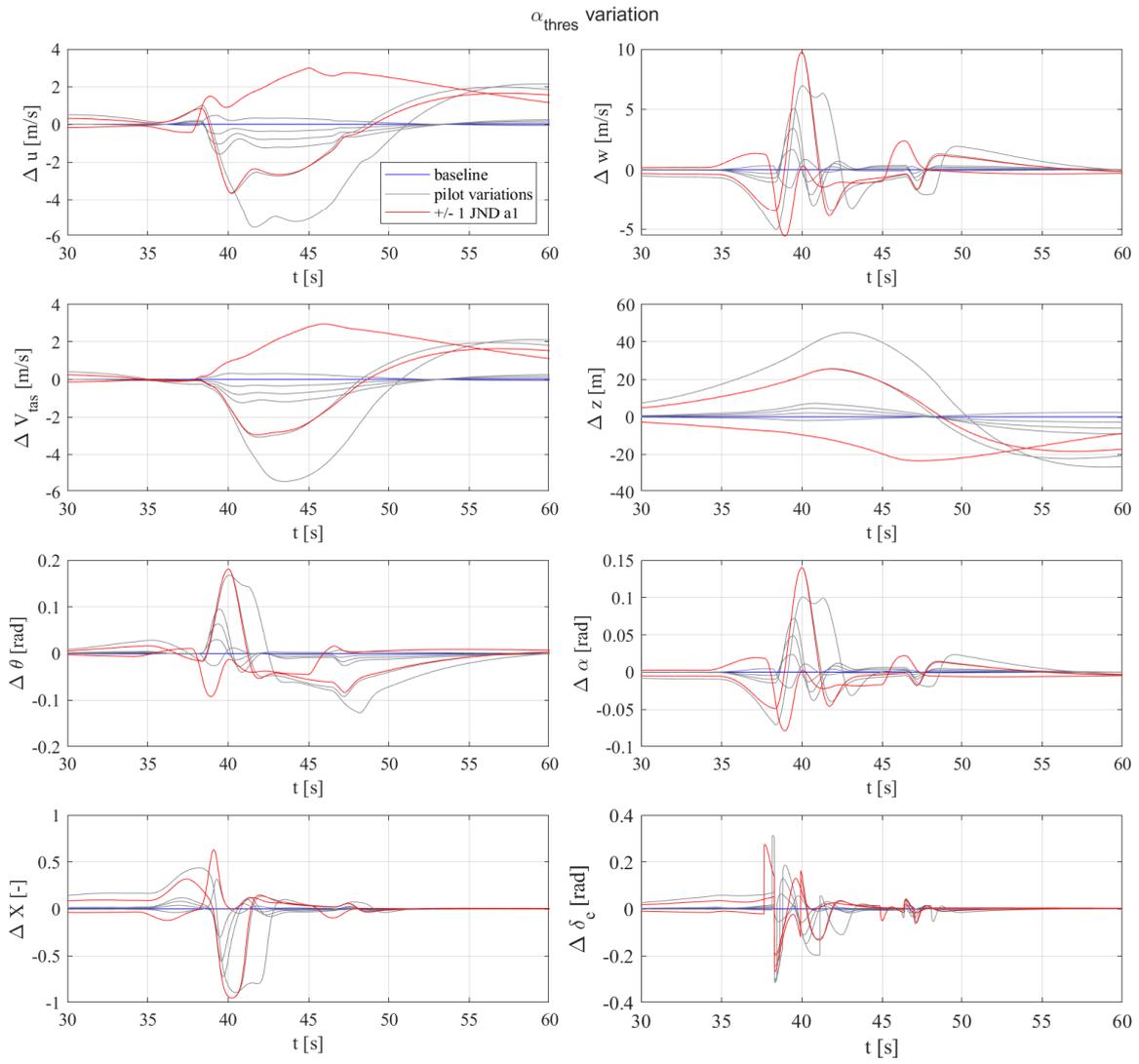


Fig. 3 Difference for relevant states between baseline model and variations in $\alpha_{\text{threshold}}$, as well as the a_1 upper and lower threshold as found by Smets et al [13].

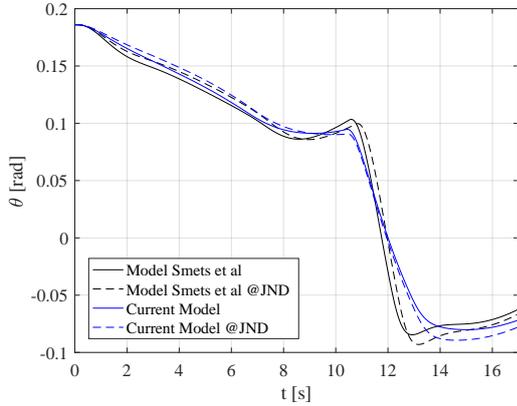


Fig. 4 Comparing pitch angle with updated C_{m_q} to the model output by [13].

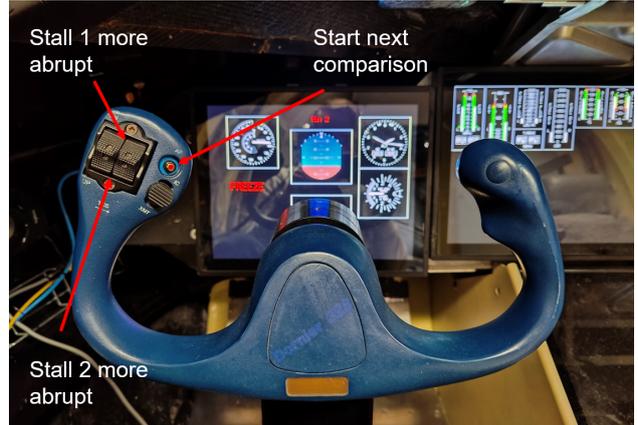


Fig. 5 The buttons used for the experiment

To what extent are the Just Noticeable Difference thresholds for a Kirchhoff stall model's stall abruptness parameter measured during a passive experiment also representative for the thresholds of pilots flying actively in a flight simulation training device?

Based on previous research [13, 15, 16], as mentioned before, two hypotheses can be formed:

H1. The active flying JND thresholds of the Kirchhoff stall model parameters will be higher than the passive thresholds. In previous research [15, 16] an active and passive threshold comparison is made, albeit that these researches investigated an absolute motion perception threshold rather than a difference perception threshold. These previous works found that the threshold in an active task were higher than in an observer task. Hence, it is expected that this will also be true for this research.

H2. The upper passive threshold for a_1 will be higher than the threshold found by Smets et al. [13], at $a_1^+ = 0.15 \pm 0.14$ written as a Weber fraction, or $a_1 = 31.96 \pm 3.86$ when written in terms of a_1 . To measure if the passive thresholds are representative for the active thresholds, the thresholds for the passive experiment must also be measured for the participants of this research. Results are expected to be comparable to the results found by Smets et al. [13]. However, because this experiment uses a different staircase procedure, which is explained in Section III.E, the 70.7% threshold will be obtained instead of the 50% threshold of Smets et al., so therefore it is expected that the average JND thresholds of the participants in this research will be higher.

B. Experiment Design

The experiment is designed as a within-subjects experiment, where all participants performed two separate experiments. The threshold for changes in stall abruptness in the flying free scenario was determined during the active experiment (see Section III.D). The passive experiment determined the passive JND threshold of the participants for changes in stall abruptness, where the autopilot would control the aircraft (see Section III.E). All odd numbered participants performed the passive experiment first, whereas all even participants performed the active experiment first. As a result, the potential learning effects of passive to active and vice versa are balanced. Both experiments were completed in the same session. The pilot was the only occupant of the simulator during the sessions. The true purpose of the experiment was disclosed to the participants at the end of the experiment, to prevent any biases from the participants confounding the experiment.

C. Procedure

The experiment measured the upper JND threshold of the stall abruptness parameter a_1 during both the active and passive experiment. The session started with a briefing, during which participants signed the consent form, filled in the demographic questionnaire, and had the opportunity to ask questions regarding the briefing and safety instructions

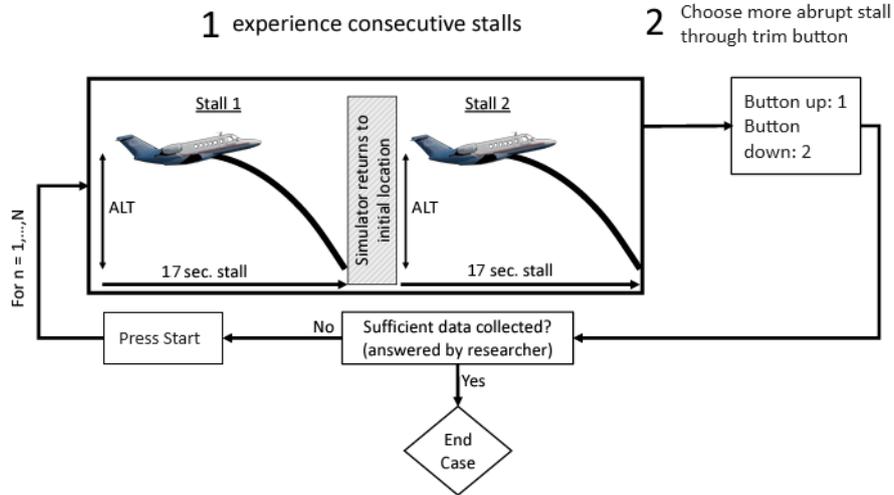


Fig. 6 Procedures of both the active and passive experiment, adapted from [13].

for SIMONA. After the briefing, the participant started with either the passive or active experiment, as mentioned in Section III.B.

The active and passive experiment consisted of several loops as shown in Figure 6. Participants experienced two consecutive stalls, with one of them having the baseline value of a_1 and the other having a higher a_1 . They were asked to identify the stall with more abrupt flow separation. This means that the task is a two alternative forced choice (2AFC) task. 2AFC tasks has an expectancy of a correct response of 50% by default, so psychometric curve goes from 50% to 100%. Therefore, the 50% threshold as found by Smets et al. [13] and the 75% threshold in this experiment are the same point on the psychometric curve. A 2AFC procedure is better than a yes/no question, such as found in the work by Smets et al. [13] and Imbrechts et al. [14] as these are prone to biases [27]. Participants were told that they would perform the loop in Figure 6 an unknown amount of times, until enough data was collected.

The procedures for the active and passive experiment are described in Section III.D and III.E, respectively. After the first experiment was completed, there was a break of approximately 15 minutes before continuing with the second part of the experiment. After the second part, there was a debrief session during which participants could comment on their experience. They were asked on which cues they based their answers, which was used for the discussion.

Throughout the experiment, participants used the trim switches to indicate which stall had more abrupt flow separation. Furthermore, they could advance to the next comparison by pushing the autopilot disconnect button. These switches can be seen in Figure 5. Neither of these buttons were used for their original purpose throughout the experiment, meaning that the pilots could not trim the aircraft. The use of these buttons was explained during the briefing. By using these buttons, the staircase was fully automated and participants could determine their own pace in the experiment.

Before starting the experiment, there was a training phase during which participants could familiarize themselves with the experiment. For both the active and passive experiment, the training consisted of a comparison between the baseline and the most extreme a_1 setting that was used in the experiment. During the active experiment training phase, extra attention was given to flying the aircraft and ensuring that the participants achieved a consistent flying performance before starting the experiment. Participants were allowed to train until they and the experimenter were confident that they could detect the differences between the baseline and a more abrupt stall.

D. Active Experiment

The active experiment used the method of constant stimuli to assess how the upper a_1 threshold as found by Smets et al. [13] translates to the active flying scenario. As this is currently unknown, the method of constant stimuli provides a first estimate such that in future research, a staircase procedure may be used to efficiently obtain a threshold for each participant. The method of constant stimuli is a rudimentary, non-adaptive method to determine a psychometric function [18]. The method uses 6-9 different, predetermined, levels of the stimuli that are presented several times to the participant. Then, by assuming a probability distribution such as the Gaussian distribution in Equation (5), the

psychometric function can be created based on the cumulative distribution function (CDF). The corresponding threshold can be found through this psychometric function, by fitting the CDF through the percentage of correct responses of the different stimuli levels. The threshold can be found by, for instance, setting $P(\varphi) = 0.75$ to get the 75% threshold [13].

$$P(\varphi) = 0.5 + 0.5 \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\varphi} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx \quad (5)$$

In order to assess how the threshold as found by Smets et al. of $a_1 = 31.9574$ translates to the active scenario, the stimuli chosen for the experiment are based on this threshold. This resulted in the test conditions as found in Table 2, where three options above and three options below the threshold are taken as stimuli, together with the threshold and the baseline value. These three options above and below the JND threshold of Smets et al. are chosen to estimate if the active threshold lies above or below the passive threshold. Each of the eight a_1 conditions is presented four times to participants, resulting in a total of 32 comparisons for the experiment.

Table 2 Active experiment settings for a_1 .

| Condition | a_1 Setting | a_1 Value |
|-----------|---------------|-------------|
| 1 | Baseline | 27.6711 |
| 2 | + 0.25 JND | 28.7427 |
| 3 | + 0.5 JND | 29.8143 |
| 4 | + 0.75 JND | 30.8858 |
| 5 | + 1.0 JND | 31.9574 |
| 6 | + 1.5 JND | 34.1005 |
| 7 | + 2.0 JND | 36.2437 |
| 8 | + 2.5 JND | 38.3869 |

To balance the experiment conditions, a Latin square is defined with all the tested a_1 settings, see Table 3. These Latin square are given labels A-H. The numbers in the top row represent the order of the conditions, so Latin square A starts with a baseline comparison, followed by a comparison with +0.25 JND. Because the active experiment tests each condition four times, four of the Latin squares A-H should be used per participant. Therefore, another Latin square is generated to balance the eight different Latin squares A-H evenly among all participants. This Latin square is split between columns 4 and 5, to generate the sequence for participants 1-8 and 9-16, see Table 4.

Table 3 Different Latin Squares with each JND value.

| Comparison → Latin square ↓ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|--------------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| A | Baseline | +0.25 JND | +0.5 JND | +1.5 JND | +0.75 JND | +1 JND | +2.5 JND | +2 JND |
| B | +0.25 JND | +1.5 JND | Baseline | +1 JND | +0.5 JND | +2 JND | +0.75 JND | +2.5 JND |
| C | +1.5 JND | +1 JND | +0.25 JND | +2 JND | Baseline | +2.5 JND | +0.5 JND | +0.75 JND |
| D | +1 JND | +2 JND | +1.5 JND | +2.5 JND | +0.25 JND | +0.75 JND | Baseline | +0.5 JND |
| E | +2 JND | +2.5 JND | +1 JND | +0.75 JND | +1.5 JND | +0.5 JND | +0.25 JND | Baseline |
| F | +2.5 JND | +0.75 JND | +2 JND | +0.5 JND | +1 JND | Baseline | +1.5 JND | +0.25 JND |
| G | +0.75 JND | +0.5 JND | +2.5 JND | Baseline | +2 JND | +0.25 JND | +1 JND | +1.5 JND |
| H | +0.5 JND | Baseline | +0.75 JND | +0.25 JND | +2.5 JND | +1.5 JND | +2 JND | +1 JND |

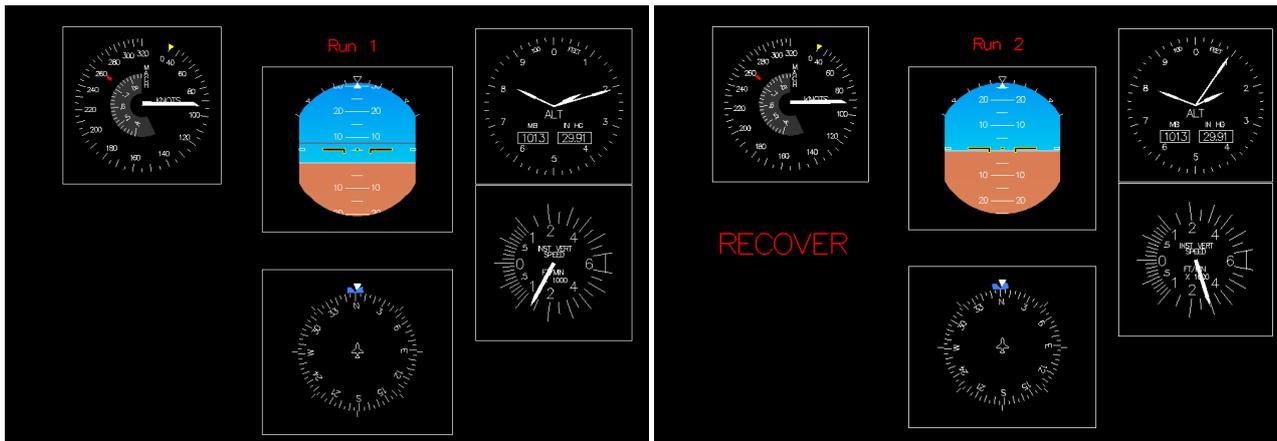
During the active experiment, participants had to follow a flight director that would guide them into the stall, as was done in the experiment by Cunningham et al. [12]. This was done as the initial simulations as shown in Section II.C indicate that differences in the approach to stall would lead to differences between the test runs that would be outside of the a_1 threshold as found by Smets et al. [13]. When participants reached full flow separation at an angle of attack of

Table 4 Used Latin squares (LS) from Table 3 for each participant.

| Subject | 1st LS | 2nd LS | 3rd LS | 4th LS | Subject | 1st LS | 2nd LS | 3rd LS | 4th LS |
|---------|--------|--------|--------|--------|---------|--------|--------|--------|--------|
| 1 | A | B | H | C | 9 | G | D | F | E |
| 2 | B | C | A | D | 10 | H | E | G | F |
| 3 | C | D | B | E | 11 | A | F | H | G |
| 4 | D | E | C | F | 12 | B | G | A | H |
| 5 | E | F | D | G | 13 | C | H | B | A |
| 6 | F | G | E | H | 14 | D | A | C | B |
| 7 | G | H | F | A | 15 | E | B | D | C |
| 8 | H | A | G | B | 16 | F | C | E | D |

$\alpha = 16^\circ$, the flight director disappeared and the "RECOVER" message as shown in Figure 7b would appear. This is the same point as when the stall autopilot would initiate the recovery.

The flight director was programmed to start at 10.65° pitch up and slowly move towards 5.925° pitch up, which matches to the trajectory flown by the stall autopilot as designed by Smets et al. [13]. This pitch up path with decreasing pitch angle ensures that the aircraft enters the stall with a 1 kts/s deceleration. Despite the decreasing pitch angle during this phase, the aircraft's vertical speed is increasing, resulting in an increasing angle of attack and therefore flow separation. Consequently, the participant will reach an angle of attack of $\alpha = 16^\circ$ within 10.5 seconds, which leads to a pitch drop of $0.45^\circ/s$.



(a) Primary Flight Display with flight director.

(b) "RECOVER" message.

Fig. 7 The primary flight display, showing the flight director for the active experiment and the recover message that is shown in both experiments.

When the "RECOVER" message appeared, an audio message saying "*Recover, recover*" was played inside the cockpit as well, ensuring that participants would start the recovery procedure immediately. During the recovery, participants were free in their handling of the aircraft. They were told to recover as they would normally do, as to not train them to perform the recovery procedure differently than their own training. It was mentioned that a consistent recovery was desired. This was one of the focus points of the training as well. Finally, participants were only able to control the elevator and throttle. The control inputs on the ailerons and rudder from the yoke and pedals were ignored in the model and controlled by the autopilot, in order to keep the stall symmetric.

E. Passive Experiment

The passive experiment used a staircase procedure to determine the upper a_1 JND threshold. There are numerous staircase procedures [17, 28–31] that have different benefits and complexity. For this research the same method is used

as in the paper by Smets et al. [13], the Parameter Estimation by Sequential Testing (PEST) procedure [17]. The PEST procedure is an adaptive staircase procedure, which optimizes the stimuli placement to quickly converge towards a threshold. By using this procedure, the found threshold can be compared to Smets et al. [13]. A few aspects of the procedure still needed to be designed:

- 1) *When to change levels* - For this a two-down, one-up (2D1U) procedure was used. This means that two correct identifications of the more abrupt stall leads to a step closer to the real parameter value, whereas one mistake leads to a step away from the real parameter value.
- 2) *What level to try next* - For this, the rules as laid out by Taylor and Creelman [17] were used. This means that the second step in a given direction had the same size as the first. At the fourth step in a given direction, the step size was doubled. And finally, at every reversal, the step size was halved. One exception in this is the first reversal, which is an exception also used by Smets et al. [13] and Imbrechts et al. [14]. To assist participants who made an early mistake from taking numerous steps to converge towards the threshold, this exception allowed them to take larger steps towards the threshold after an early mistake.
- 3) *When to stop* - The procedure was stopped if one of the following three criteria was met. The procedure stopped if the step size was equal to or smaller than $1/64^{\text{th}}$ of the original step size, if the number of reversals was eight, or if 32 comparisons were made. The maximum of 32 comparisons was chosen because it is the amount of trials for the active experiment as well.

For this experiment, the starting level was $a_1 = 50$ with a step size of 7.5. By setting the step size to 7.5, the participant reached 27.5 after 3 steps towards the threshold, which was set to the baseline value of 27.6711. This prevented participants from reaching four consecutive correct answers before the first reversal, and avoided an early doubling of the step size and lead to an efficient staircase procedure.

The procedures of the passive experiment were the same as in the active experiment, except for the flight director as shown in Figure 7a, which was removed for the passive experiment. The "RECOVER" message remained, as well as the audio message, to give participants a sense of when the autopilot initiated the recovery.

F. Apparatus

The experiment was performed in the SIMONA Research Simulator at the Faculty of Aerospace Engineering of Delft University of Technology, see Figure 8. SIMONA is a 6-degrees of freedom hexapod simulator. Participants were seated in the captain's seat, shown on the left in Figure 9, where a control column is present. This is because the Cessna Citation II is also equipped with a control column. Participants wore a noise canceling headset to mask any noise coming from the motion system. In the background, static engine noise played on the cockpit speakers throughout the experiment, to further mask the noise from the motion system.



Fig. 8 The SIMONA research simulator [32].



Fig. 9 Cockpit view from inside SIMONA.

During the experiment, participants could observe the crucial flight states on the primary flight display, which showed multiple instruments as can be seen in Figure 7a. Furthermore, they had a second display that showed the engine parameters as shown in Figure 10. Furthermore, participants were provided with an outside visual covering $180^\circ \times 40^\circ$, which is provided by three projectors that have a resolution of 1280×1024 pixels and a refresh rate of 60 Hz [32]. The outside visual image was provided by a FlightGear visual database.

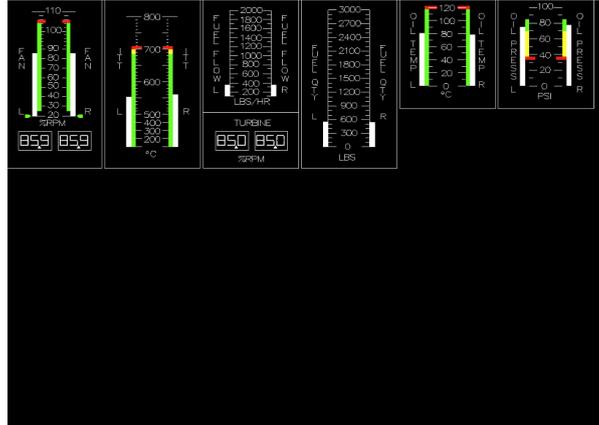


Fig. 10 Engine display as used in the experiment [14].

The settings for the washout filter used in the experiment are found in Table 5. These settings were mostly similar to the settings used by Smets et al. [13] and Imbrechts et al. [14], except for the pitch gain which was 0.7 instead of 0.5 in the work by Imbrechts et al. [14]. The roll, yaw, and sway motions were not used for this experiment, since only symmetrical stalls were simulated.

Table 5 Motion Filter Settings.

| | | High-pass filters | | | | Low-pass filters | |
|----------------|-----------|-------------------|-----------|----------------|-----------|------------------|-----------|
| ω_{n_q} | 1.0 rad/s | ω_{n_z} | 2.0 rad/s | ω_{n_x} | 1.2 rad/s | ω_{n_x} | 2.4 rad/s |
| ζ_q | 0.7 | ζ_z | 0.7 | ζ_x | 0.7 | ζ_x | 0.7 |
| ω_{b_q} | 0.0 rad/s | ω_{b_z} | 0.3 rad/s | ω_{b_x} | 0.0 rad/s | | |
| K_q | 0.7 | K_z | 0.5 | K_x | 0.5 | | |

G. Participants

A total of 16 participants participated in the experiment, all of whom had experience flying twin-turbine aircraft. The participants had a mean age of 48.5 years (standard deviation (SD) = 9.1 years), and a mean number of flight hours of 9,567 hours (SD = 6,961 hours). The group consisted of 10 Captains, 4 First Officers, and 2 Second Officers. Four participants had a Cessna Citation II type rating, eight had a Boeing 777/787 type rating, four had a Boeing 737 type rating, one participant had a Embraer E175, E190, and E195-E2 type rating, one participant had a Boeing 747 rating, and one participant had a Gulfstream G650 type rating. There were several participants who had multiple ratings.

Before the experiment started, it was indicated to the participants that their stall recovery performance was not evaluated and they were asked to focus on detecting differences in the model's response. Pilots voluntarily participated in this experiment and gave informed consent before starting. This research was approved by the Human Research Ethics Committee of TU Delft under application number 3643.

H. Data Analysis

1. Passive experiment

For the passive experiment, the final JND threshold of a participant was determined by averaging the staircase's a_1 values across the last three reversals, as was done by Smets et al. [13]. The resulting threshold is denoted as a_1^+ , as it is the upper threshold for a_1 . From the individual thresholds, the average threshold of the entire group was determined. Furthermore, the individual thresholds of each participant from Smets et al. [13] were compared to the data set of this experiment through the Mann-Whitney U test, since neither data set is normally distributed, to see if the results are significantly different.

The a_1^+ threshold that was found in this experiment is the 70.7% threshold because of the 2U1D staircase used [29], as opposed to a 50% threshold for a one-up, one-down procedure used by Smets et al. [13]. This 70.7% threshold is valid if the psychometric curve for percentage of correctness is from 0% to 100%. However, due to 2AFC design, this curve is from 50% to 100%, since the default percentage of a correct response for a two-choice experiment is 50%. Consequently, the 70.7% threshold that resulted from the 2U1D staircase is shifted to the 85.35% threshold point on the 2AFC psychometric function.

Next to this, there is a second method to determine the threshold of the participant group, which is less influenced by outliers than taking the average. This was done by fitting the psychometric function through the average percentage of correct responses of each individual a_1 value used in the experiment based on the combined answers of all participants. This method was also used by Smets et al. [13] and the resulting μ of both CDFs can therefore be directly compared. Because this experiment used the 2AFC question, the μ for this experiment represents 75% threshold, whereas the μ of Smets et al. [13] represents the 50% point on the CDF. However, both are the same point on the psychometric function and can therefore be compared.

2. Active experiment

The method of constant stimuli used for the active experiment, which does not give a direct threshold in its current set up [18]. Therefore, only the combined percentages of correct responses per participant for each of the eight test conditions were taken as the data set and used to estimate the JND threshold of the participant group. This was done by fitting the CDF through the average of correct responses, which resulted in the psychometric function for the active experiment. Then, the 75% threshold could also be determined by setting $P(\varphi) = 0.75$ in Equation (5) to estimate the corresponding active JND threshold.

To support this found threshold, a pairwise comparison between the data sets of each a_1 condition was done. First, the data was tested for normality with the Shapiro-Wilk test. Following this, a Friedmann ANOVA was used to determine if the different test conditions are significantly different. Then, the Wilcoxon signed rank test provided a pairwise comparison to highlight which a_1 conditions are significantly different from each other. By comparing this to the distribution found at the baseline comparison, the conditions that are significantly different from the default chance were identified.

IV. Results

The results of the passive and active experiment are discussed separately. An example of the results of one of the participants can be seen in Figure 11.

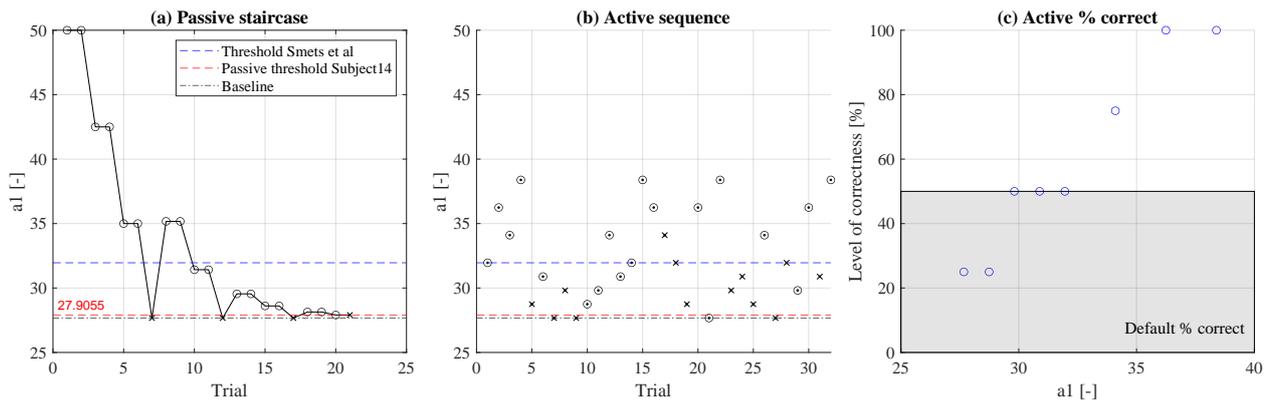


Fig. 11 Example results of both the passive and active experiment for Subject 14, presented together with the percentages of correct answers for the active experiment.

A. Passive experiment

The staircase used in this experiment yielded the 70.7% threshold for the passive experiment. The found threshold was based on the average of the last three reversals for each participant, which can be seen in Figure 11a. The resulting thresholds for all participants can be seen in Figure 12. The average threshold, with the 95% confidence interval, is $a_1^+ = 30.62 \pm 2.62$ [-]. When shown as a Weber fraction ($\frac{\Delta I}{I}$) with the 95% confidence interval, this becomes: $a_1^+ = 0.11 \pm 0.094$. The median threshold is at $a_1 = 29.16$ [-].

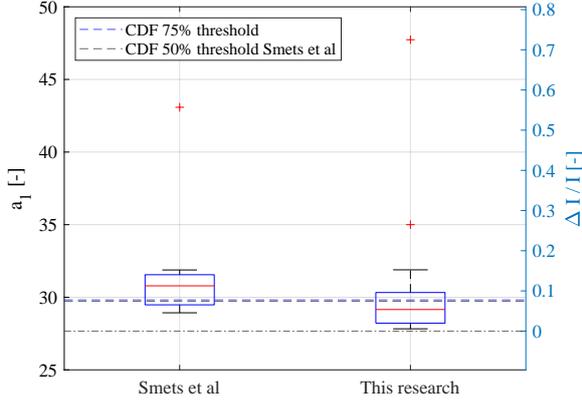


Fig. 12 Boxplot representations of the passive staircase results as well as the CDF thresholds from the work by Smets et al [13] and this research.

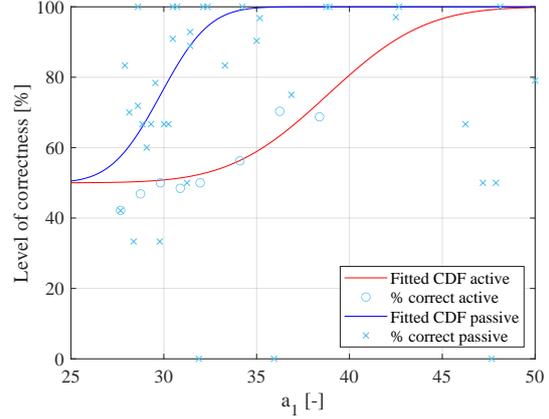


Fig. 13 CDF of the active and passive experiment, fitted through the percentage correct for each tested a_1 value.

The data of the entire staircase procedure was also used to construct the psychometric function for the passive threshold. The CDF as shown in Equation (5) was fitted through the average percentage of correct responses of all participants for each a_1 level, which resulted in the blue CDF in Figure 13. From this CDF, an estimate of the 75% threshold was made, which can be compared to the CDF threshold found by Smets et al. [13]. The CDF of this experiment gave a 75% threshold of $P(75\%) = 29.82$ or $a_1^+ = 0.078$ when expressed as a Weber fraction. This is similar to the threshold of Smets et al. which is $P(50\%) = 29.72$, or $a_1^+ = 0.074$.

There was one participant who changed his cues for detecting the more abrupt flow separation during the passive experiment. He started the staircase by going down, however, after the first 5 comparisons, the steps were taken back up where a consistent incorrect answer was given at $a_1 = 50$ [-]. After 7 runs at $a_1 = 50$ [-] with a consistent wrong answer, the experimenter intervened and asked if perhaps the participant switched his cues. This was confirmed by the participant and consequently, the experimenter told the participant that he coupled the cues to the wrong answer. After this, the participant continued the staircase as normal. The 8 consecutive wrong answers at $a_1 = 50$ [-] were omitted from the data to fit the CDF. This resulted in the CDF as shown in Figure 13. Omitting this data had little influence on the outcome of the experiment, as the current CDF uses $\mu = 29.8188$, and $\sigma = 2.0769$, whereas including these measurements resulted in $\mu = 29.8187$, and $\sigma = 2.0772$ as well.

B. Active experiment

1. Statistical analysis

The percentage of correctness for each condition of every participant was taken as a data set and each of the data sets was tested for normality. Based on the Shapiro-Wilk test, it was concluded that only the data for conditions 3 and 5 (+0.5 JND and +1 JND respectively) was normally distributed. Therefore, a Friedmann ANOVA was used to determine if there is a significant difference in the data. It was found that there is a statistically significant difference in the correctness levels for the a_1 conditions ($\chi^2(7) = 16.105$, $p = 0.024$).

Following this, a pairwise comparison between the different conditions is done through a Wilcoxon signed rank test, which is given in Table 6. From this, it was concluded that there is not a significant difference between Conditions 1-6. However, Conditions 7 shows a significant difference with Conditions 1-5. Condition 8 shows a significant difference

with Conditions 1 and 5, and generally lower p-values that suggest differences with Conditions 1-6, although no statistically conclusive statements can be made about it. However, from this pairwise comparison, it can be concluded that Conditions 1-6 do not significantly differ, meaning that the data in these conditions are mostly due to chance. This comes from the fact that the distribution for Condition 1 is due to chance, as it is the baseline comparison. The fact that Conditions 7 and 8 are significantly different from Condition 1 means that these distributions are not due to chance and are therefore the first conditions significantly higher than 50% on the psychometric curve.

Table 6 Pairwise comparison significance from Wilcoxon signed rank test with statistically significant differences highlighted.

| | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|-------|-------|-------|-------|-------|--------------|--------------|
| 1 | 0.559 | 0.257 | 0.113 | 0.516 | 0.073 | 0.005 | 0.017 |
| 2 | | 0.775 | 0.677 | 0.981 | 0.441 | 0.030 | 0.078 |
| 3 | | | 1.000 | 0.807 | 0.636 | 0.020 | 0.156 |
| 4 | | | | 0.844 | 0.613 | 0.010 | 0.063 |
| 5 | | | | | 0.560 | 0.012 | 0.022 |
| 6 | | | | | | 0.053 | 0.128 |
| 7 | | | | | | | 0.766 |

To see the position of Conditions 7 and 8 on the psychometric curve, the CDF in Equation (5) was fitted through the active data, which can be seen in Figure 13 in red. From this psychometric function, the 75% threshold was estimated as well, which is $P(75\%) = 38.83$ or $a_1^+ = 0.40$ when expressed as a Weber fraction. This is significantly higher than the passive JND threshold of $P(75\%) = 29.82$ or $a_1^+ = 0.078$ found in this research and therefore confirms hypothesis **H1**.

2. Analysis of participants' control consistency

An analysis on each of the participants' consistency in control behavior was performed. This is done by calculating the average control input for every baseline run at each point in time. Then, the root mean square (RMS) of the difference between each run and the average is calculated. The average RMS is taken as a metric to compare between the participants. In Figure 14, the control inputs from Subject 14 and resulting pitch angle and flow separation point can be observed, which show consistent control inputs throughout the experiment. Figure 11b and c also shows that Subject 14 correctly identified the more abrupt flow separation of the highest a_1 conditions.

Subject 15, on the other hand, was less consistent with his control inputs, as can be seen in Figure 15. This participant could not correctly identify the highest a_1 conditions in this experiment, as shown in Figure 16b and c. The results for all participants are summarized in Table 7. The clear outlier in this data is Subject 12, who had a consistent control input and lower percentages correct. This participant indicated afterwards that he switched strategies halfway through the active experiment. Despite this, his data is included in all of the analyses. This verified the comments made by participants, that their own control inputs influenced their ability to detect the differences.

Table 7 Analyzed control behavior consistency expressed in Root Mean Square (RMS) of the average elevator inputs δ_e and the percentage of correct answers for condition 7 and 8 for each participant.

| Subject | RMS [rad] | % correct #7 | % correct #8 | Subject | RMS [rad] | % correct #7 | % correct #8 |
|---------|-----------|--------------|--------------|---------|-----------|--------------|--------------|
| 1 | 0.0279 | 75 | 25 | 9 | 0.0187 | 75 | 75 |
| 2 | 0.0335 | 50 | 75 | 10 | 0.0212 | 50 | 25 |
| 3 | 0.0211 | 75 | 75 | 11 | 0.0172 | 50 | 50 |
| 4 | 0.0224 | 75 | 100 | 12 | 0.0177 | 25 | 25 |
| 5 | 0.0228 | 75 | 100 | 13 | 0.0192 | 75 | 100 |
| 6 | 0.0187 | 100 | 100 | 14 | 0.0154 | 100 | 100 |
| 7 | 0.0188 | 100 | 75 | 15 | 0.0265 | 75 | 50 |
| 8 | 0.0367 | 75 | 75 | 16 | 0.0228 | 50 | 50 |

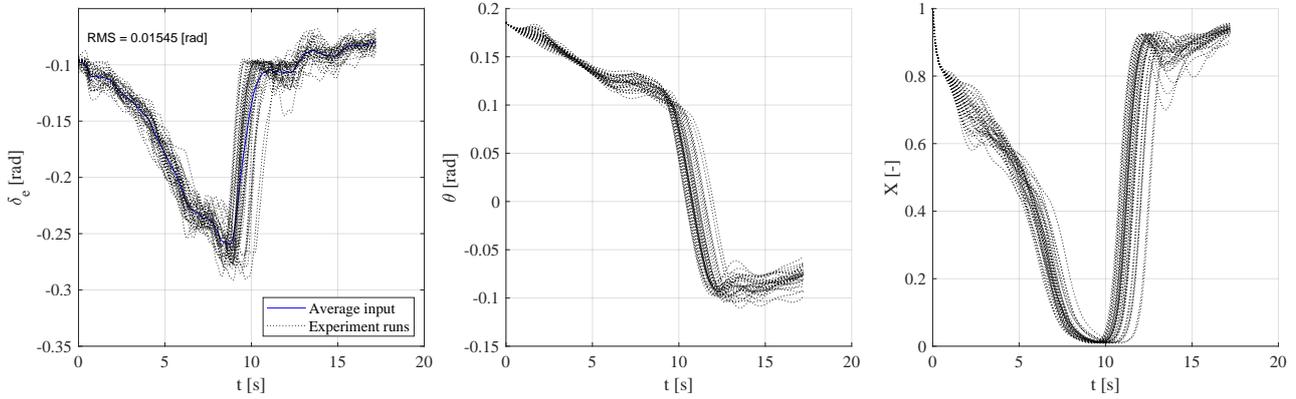


Fig. 14 Elevator inputs for all baseline runs from Subject 14, together with the pitch angle and flow separation point.

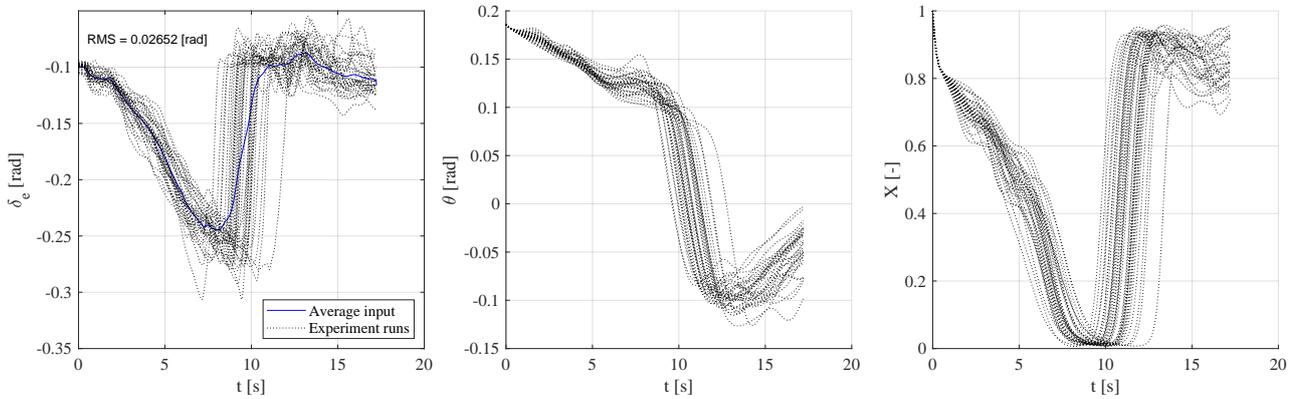


Fig. 15 Elevator inputs for all baseline runs from Subject 15, together with the pitch angle and flow separation point.

3. Analysis of comments by participants

Apart from the data analysis, the comments participants gave after completing the experiment were analyzed. All participants indicated that they had more difficulty in the active experiment to detect the differences. They explained that they felt that they had less mental capacity to focus on detecting the differences as they were busy controlling the aircraft. Furthermore, they had the idea that their control input was influencing the results of the stall abruptness and therefore doubted whether or not the difference they felt was due to their inputs or due to the differences in the model. Finally, participants noted, when learning the true intent of the experiment, that they would agree that each of the conditions presented to them in the active experiment could be used for stall training.

V. Discussion

The CDFs as shown in Figure 13 provide insights into how the threshold as found in the passive experiment translates to the active experiment. The 75% thresholds for both CDFs lie at $P(75\%) = 29.82$ and $P(75\%) = 38.83$ for the passive and active experiment, respectively. When written as Weber fractions, the thresholds are 0.078 and 0.40 for the passive and active experiment, respectively, indicating that the threshold for the active experiment is over five times higher. On top of that, the entire psychometric function of the active experiment lies to the right of the passive psychometric function, as shown in Figure 13. Consequently, hypothesis **H1** is accepted. This means that the sensitivity of participants for changes in stall abruptness for pilots flying a stall themselves is decreased.

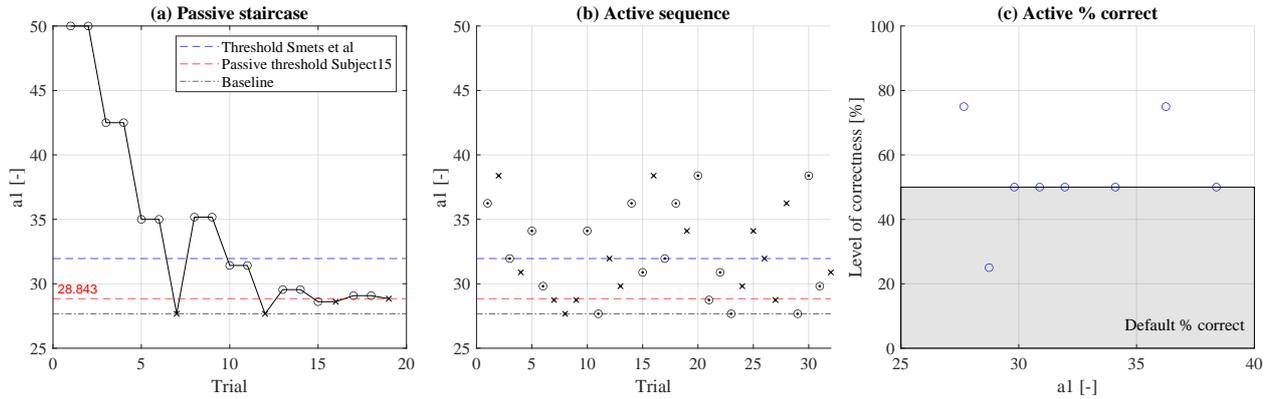


Fig. 16 Results of both the passive and active experiment for Subject 15, presented together with the percentages correct for the active experiment.

To compare the results of the passive experiment with the data as found by Smets et al. [13], a Mann-Whitney U test was performed to verify that the individual thresholds found are not significantly different. This test resulted in $U = 35$, with significance $p = 0.076$, which means that there is not a significant difference between the two data sets. It was hypothesized that the staircase procedure of this research would lead to a higher JND threshold, since it results in the 70.7% threshold as opposed to the 50% threshold resulting from the staircase procedure used by Smets et al. [13]. However, from the average JND threshold visible in the boxplots in Figure 12, it can be concluded that the JND threshold for this experiment is not higher than the JND threshold found by Smets et al. [13]. The average threshold of this experiment is $a_1^\dagger = 0.11 \pm 0.094$, whereas the threshold found by Smets et al. [13] is $a_1^\dagger = 0.16 \pm 0.14$. Therefore, hypothesis **H2** is rejected.

There are two possible explanations as to why the average passive JND threshold found in this research is lower than the threshold found in Smets et al. [13]. Firstly, Smets et al. asked their participants if they noticed a difference between two test runs. For their participants, this meant that there could be differences possible in all aspects of the simulation. This experiment asked participants to identify the run with the most abrupt flow separation, which is a much clearer task. The Yes/No questions of Smets et al. is particularly prone to biases, as participants may be more inclined to answering yes, although they did not feel any difference or were unsure if there was a difference [27]. Their solution was to have every third test run compare two baseline runs with each other, which should lead to participants indicating that they did not feel a difference. As a result, a third of their measurements was dedicated to detecting the bias, whereas this research used every measurement to converge towards the threshold.

How often participants answered that they felt a difference in these baseline comparisons, allowed Smets et al. to assess the participants' consistency and reliability. This is the second reason that the procedure used by Smets et al. is less effective than the procedure of this experiment, as they may have confused the participants more by making each third run a baseline versus baseline comparison. This bias and confusion becomes evident when comparing the psychometric function of Smets et al. [13] to the psychometric function resulting from this research. Smets et al. reported a psychometric function mean of $\mu = 29.72$ with a variance of $\sigma = 2.33$, whereas the current research yielded a mean of $\mu = 29.82$ with a variance of $\sigma = 2.08$. The higher average of the individual staircase thresholds is indicative of a less efficient staircase method that, despite these inefficiencies, still gives a similar result in the mean of the psychometric curve.

Using the method of constant stimuli for the active experiment might in hindsight have been a confusing factor for the current experiment, which may have contribute to the higher active JND threshold. One of the reasons is that, since participants randomly started with one of the Latin squares with active conditions, there was no direct comparison from the training to the real experiment. In the training, participants trained with detecting differences between the most extreme a_1 values they would encounter, which are $a_1 = 50$ [-] and $a_1 = 38.3869$ [-] for the passive and active experiment, respectively. For the passive experiment, the first runs also contained the comparison with $a_1 = 50$ [-], so they could directly recognize the cues from the training. For the active experiment, this was not necessarily true. For instance Subject 1 and 11 and Subject 8 and 10 started with Latin square A and H, respectively. Both Latin squares started with relatively low a_1 values in the first three or four runs. These subjects did, therefore, not recognize any of the

cues that they experienced during the training, which can be solved by using a staircase procedure.

This confusion became evident from the responses of the participants during the debriefing. Most participants indicated that they felt that they guessed for most of the experiment and were unsure about their performance for the active part of the experiment. Another contributing factor is that the highest a_1 value used in the experiment is $a_1 = 38.3869$, whereas the resulting 75% threshold of the combined answers is $P(75\%) = 38.83$. The used conditions were, therefore, insufficient to assess the entire psychometric function for the active task, since the entire CDF curve is based on half of the threshold information. Therefore, for future research, it is better to assess the threshold for active flying including also higher values of a_1 .

This can, for instance, be done by a staircase procedure as well. As seen in the results of the passive experiment, the PEST procedure that was designed placed the stimuli efficiently to obtain an accurate estimate of the threshold with a minimum amount of trials. Every participant finished the experiment in under 32 trials, except for Subject 2 who was discussed separately in subsection IV.A. In addition, the psychometric function was fit with a much lower variance than for instance done by Smets et al. [13].

Unfortunately, due to the fact that the highest a_1 condition was not sufficient to reach the 100% correctness level for the active experiment, the analysis of the control behavior cannot lead to conclusive answer on what the influence of control behavior of the participants is on the resulting percentages correct. Furthermore, because there are only 4 test points per a_1 condition, the individual percentages correct of the participants are too much susceptible to chance to compare to the control consistency. However, it is an interesting aspect to investigate in future research. For instance, when an individual threshold is found, the consistency in control behavior can be directly compared to verify whether or not there is an influence on the obtained thresholds.

One of the points the participants raised about the simulation, was its odd entry into stall. The simulation from Smets et al. [13], which was used in this research as well, started at 10° pitch up and lowered the pitch attitude to about 5° to reach a 1 kts/s deceleration into the stall. Participants mentioned that their usual stall training included a pitch up behavior due to, for instance, the altitude hold mode of the autopilot, a scenario used by Stepanyan et al. [33] and Lombaerts et al. [34] in their stall model research. This creates a more realistic scenario for the participants.

Finally, a shortcoming of the simulation was that there is currently no force feedback implemented. The control column in SIMONA was configured as a passive mass-spring damper system. As a result, participants could not feel the elevator feedback as they would normally have in the real Cessna Citation II. Furthermore, all participants fly the Cessna Citation II, Boeing 737, 777/787, or the Embraer family, all of which have a control column with force feedback. In future research, implementing force feedback may contribute to a more complete feel of the aircraft. Some of the participants indicated that they already noticed a difference in the amount of back-pressure they needed to apply between the test runs, so adding accurate force feedback to the simulation is an essential next step.

The results from this research can be expanded to the JND thresholds for other stall model parameters as found by Smets et al. [13]. The results suggest that the lower a_1 JND threshold, as well as the upper and lower τ_1 would likely be further from the baseline in an active flying scenario, than the passive thresholds found by Smets et al..

Finally, the found sensitivity of the participants can be related to the uncertainty of the model as created by Van Ingen et al. [20]. They found the baseline value for $a_1 = 27.6711$ with standard deviation $\sigma = 6.72$ or 0.248 when shown as a fraction. Brill [24] used a slice-based modeling approach to identify the a_1 value for the Cessna Citation II aircraft and found a value of $a_1 = 34.1856$, which is a difference of 23.5% as compared to the value used in this research. Both these uncertainties lie within the found active JND threshold of $a_1^+ = 0.40$ found by this research. This implies that the modeling techniques as used by Van Ingen et al. [20] and Brill [24] provide a model that is accurate enough for simulator-based stall training.

The results of this research show that there is a decreased sensitivity for changes in stall abruptness when pilots are in active control of the stall model. Furthermore, it gives a threshold where participants are able to perceive the differences in the model, which can be used by the authorities to converge towards a required accuracy for stall models to be used in simulators. As a result, the "*within the realms of confidence*" [9] as mentioned by EASA can be transformed towards an actual required accuracy, which can be the next step to ensuring that simulator-based stall training is truly effective due to models that are accurate enough to facilitate this.

VI. Conclusions

This research provides, for the first time, a quantitative estimate on how the JND thresholds for changes in stall abruptness for symmetric stall in a passive, observer scenario translate to an active pilot-in-command scenario. It was found that the passive JND threshold is 0.11 ± 0.094 , expressed as a Weber fraction. This JND threshold is lower than the

JND threshold found by previous research due to the use of an improved staircase procedure. From the psychometric function fitted through all answers of the participants, a 75% threshold of $P(75\%) = 29.82$ or 0.078 expressed as Weber fraction is found. Furthermore, this research successfully found an active JND threshold for stall abruptness, which lies at $P(75\%) = 38.83$ or 0.40. The active threshold is therefore over five times higher than the passive threshold. Furthermore, the entire psychometric curve of the active experiment lies to the right of the passive curve, indicating that pilots have a decreased sensitivity for changes in stall abruptness when flying themselves. The resulting JND thresholds of this research provide a basis for regulatory bodies to define and implement more precise accuracy standards for stall models in simulators, thereby enhancing the fidelity and effectiveness of pilot stall training.

Acknowledgments

The author would like to express his gratitude to all the pilots who voluntarily participated in this research. Next to them, a sincere thank you to the daily supervisors of this master thesis research, dr.ir. D.M. Pool and dr.ir. C.C. de Visser. Furthermore, the author would like to thank the American Institute of Aeronautics and Astronautics for awarding him the AIAA Luis de Florez Graduate Award for the research proposal of this research. Finally, a thank you to the technicians and responsible personnel of the SIMONA Research Simulator at the Faculty of Aerospace Engineering.

References

- [1] IATA, "Loss of Control In-Flight Accident Analysis Report," Tech. rep., International Air Transport Association, 2019.
- [2] "ICAO Doc 10011: Manual on Aeroplane Upset Prevention and Recovery Training," Tech. rep., International Civil Aviation Organization, 2014.
- [3] NTSB, "Aircraft Accident Report; Loss of Control on Approach; Colgan Air, Inc," Tech. rep., National Transportation Safety Board, Washington, DC, 2 2010.
- [4] OVV, "Turkish Airlines Flight TK1951, a Boeing 737-800, Crashed during approach," Tech. rep., Dutch Safety Board, 2010. URL www.safetyboard.nl.
- [5] d'Enquêtes et d'Analyses pour la sécurité de l'aviation civile, B., "Final Report on AF447," Tech. rep., Ministère de l'Écologie, du Développement durable, des Transports et du Logement, 2009. URL www.bea.aero.
- [6] "Loss of control prevention and recovery training," <https://www.easa.europa.eu/en/downloads/71682/en>, feb 2019.
- [7] "Upset Prevention and Recovery Training," https://www.faa.gov/documentLibrary/media/Advisory_Circular/AC_120-111_CHG_1.pdf, jan 2017.
- [8] "Part 60 - Flight Simulation Training Device Initial and Continuing Qualification and Use," Tech. rep., Federal Aviation Administration, United States Department of Transportation, 10 2021. URL <https://www.ecfr.gov/current/title-14/part-60>.
- [9] EASA, "Certification Specifications for Aeroplane Flight Simulation Training Devices 'CS-FSTD(A)'," Tech. rep., European Aviation Safety Agency, 2018.
- [10] Schroeder, J. A., Burki-Cohen, J. S., Shikany, D., Gingras, D. R., and Desrochers, P. P., "An evaluation of several stall models for commercial transport training," *AIAA Modeling and Simulation Technologies Conference*, 2014, p. 1002.
- [11] Grant, P. R., Moszczynski, G. J., and Schroeder, J. A., "Post-stall flight model fidelity effects on full stall recovery training," *2018 Modeling and Simulation Technologies Conference*, American Institute of Aeronautics and Astronautics Inc, AIAA, 2018. <https://doi.org/10.2514/6.2018-2937>.
- [12] Cunningham, K., Shah, G. H., Murphy, P. C., Hill, M. A., and Pickering, B. P., "Pilot sensitivity to simulator flight dynamics model formulation for stall training," *AIAA Scitech 2019 Forum*, American Institute of Aeronautics and Astronautics Inc, AIAA, 2019. <https://doi.org/10.2514/6.2019-0717>.
- [13] Smets, S., De Visser, C. C., and Pool, D. M., "Subjective Noticeability of Variations in Quasi-Steady Aerodynamic Stall Dynamics," *AIAA Scitech 2019 Forum*, 2019. <https://doi.org/10.2514/6.2019-1485>.
- [14] Imbrechts, A., De Visser, C. C., and Pool, D. M., "Just Noticeable Differences for Variations in Quasi-Steady Stall Buffet Model Parameters," *AIAA Scitech 2022 Forum*, 2022. <https://doi.org/10.2514/6.2022-0510>.

- [15] Hosman, R. J. A. W., and van der Vaart, J. C., “Vestibular Models and Thresholds of Motion Perception. Results of tests in a flight simulator,” Tech. rep., Delft University of Technology, Delft, 1978. URL <http://resolver.tudelft.nl/uuid:72bb1e55-7304-459f-a47c-dae7984418e3>, IR-265.
- [16] Valente Pais, A. R., Pool, D. M., De Vroome, A. M., Van Paassen, M. M., and Mulder, M., “Pitch motion perception thresholds during passive and active tasks,” *Journal of Guidance, Control, and Dynamics*, Vol. 35, No. 3, 2012, pp. 904–918. <https://doi.org/10.2514/1.54987>.
- [17] Taylor, M. M., and Creelman, C. D., “PEST: Efficient estimates on probability functions,” *The Journal of the Acoustical Society of America*, Vol. 41, No. 4A, 1967, pp. 782–787.
- [18] Simpson, W. A., “The method of constant stimuli is efficient,” *Perception & psychophysics*, Vol. 44, 1988, pp. 433–436.
- [19] Van Horssen, L., De Visser, C. C., and Pool, D. M., “Aerodynamic Stall and Buffet Modeling for the Cessna Citation II Based on Flight Test Data,” *2018 AIAA Modeling and Simulation Technologies Conference*, 2018, p. 1167. <https://doi.org/10.2514/6.2018-1167>.
- [20] Van Ingen, J., Visser, C. C., and Pool, D. M., “Stall Model Identification of a Cessna Citation II from Flight Test Data Using Orthogonal Model Structure Selection,” *AIAA Scitech 2021 Forum*, 2021, p. 1725. <https://doi.org/10.2514/6.2021-1725>.
- [21] Fischenberg, D., “Identification of an Unsteady Aerodynamic Stall Model from Flight Test Data,” *20th Atmospheric Flight Mechanics Conference*, American Institute of Aeronautics and Astronautics Inc, 1995, pp. 138–146.
- [22] Van Horssen, L. J., “Aerodynamic Stall Modeling for the Cessna Citation II,” Master’s thesis, Delft University of Technology, Delft, 2016. URL <https://repository.tudelft.nl/>.
- [23] Smets, S. C. E., “Subjective Noticeability of Variations in Quasi-Steady Aerodynamic Stall Dynamics,” Master’s thesis, Delft University of Technology, Delft, 2017. URL <https://repository.tudelft.nl/>.
- [24] Brill, P. A. R., “Improving Stall Model Accuracy through Optimal Data Slicing by Analyzing Kirchhoff Stall Parameter Estimate Behaviour,” Master’s thesis, Delft University of Technology, Delft, 2023. URL <https://repository.tudelft.nl/>.
- [25] Mulder, J., van Staveren, W., van der Vaart, J., de Weerd, E., de Visser, C., in ’t Veld, A., and Mooij, E., “Flight Dynamics,” Faculty of Aerospace Engineering, Delft University of Technology, 3 2013. Lecture Notes AE3212.
- [26] Van den Hoek, M., de Visser, C., and Pool, D., “Identification of a Cessna Citation II Model Based on Flight Test Data,” *Fourth CEAS Specialist Conference on Guidance, Navigation and Control*, Springer, 2017, pp. 259–277.
- [27] Kingdom, F. A., and Prins, N., *Psychophysics: A practical introduction*, 2nd ed., Mica Haley, 2016.
- [28] Cornsweet, T. N., “The Staircase-Method in Psychophysics,” *The American Journal of Psychology*, Vol. 75, No. 3, 1962, pp. 485–491. URL <http://www.jstor.org/stable/1419876>.
- [29] Levitt, H., “Transformed up-down methods in psychoacoustics,” *The Journal of the Acoustical society of America*, Vol. 49, No. 2B, 1971, pp. 467–477.
- [30] Leek, M. R., “Adaptive procedures in psychophysical research,” *Perception & Psychophysics*, Vol. 63, No. 8, 2001, pp. 1279–1292.
- [31] Pentland, A., “Maximum likelihood estimation: The best PEST,” *Perception & Psychophysics*, Vol. 28, No. 4, 1980, pp. 377–379.
- [32] Stroosma, O., Van Paassen, M., and Mulder, M., “Using the SIMONA Research Simulator for human-machine interaction research,” *AIAA modeling and simulation technologies conference and exhibit*, 2003, p. 5525.
- [33] Stepanyan, V., Krishnakumar, K. S., Kaneshige, J., and Acosta, D. M., “Stall recovery guidance algorithms based on constrained control approaches,” *AIAA Guidance, Navigation, and Control Conference*, 2016, p. 0878.
- [34] Lombaerts, T., Schuet, S., Stepanyan, V., Kaneshige, J., Hardy, G., Shish, K., and Robinson, P., “Piloted simulator evaluation results of flight physics based stall recovery guidance,” *AIAA Guidance, Navigation, and Control Conference, 2018*, American Institute of Aeronautics and Astronautics Inc, AIAA, 2018. <https://doi.org/10.2514/6.2018-0383>.

Part II

Preliminary Thesis Report

Already Graded for AE4020

Background Information on Stall

This chapter provides a literature review on stall and stall modeling. First, the stall phenomenon itself will be explained, followed by a description on how to capture stall in a model. Finally, the key takeaways for this research are summarized in the conclusions.

2-1 Aircraft Stall

2-1-1 Aerodynamic Stall

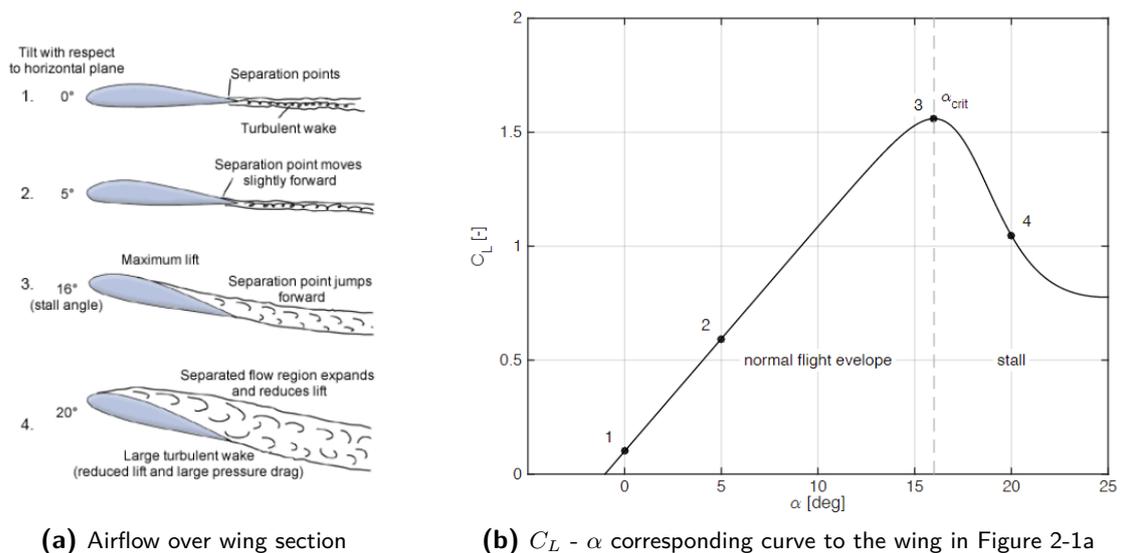


Figure 2-1: Progress of trailing edge flow separation [12]

Aerodynamic stall happens when the airflow over a wing separates from the wing. Flow separation can happen for several reasons. One of them is when the separation point of the

airflow gradually moves forward with increasing angle of attack, which is called trailing edge stall. This can be seen in Figure 2-1. However, at some point, called the critical angle of attack or the stall angle of attack, the wing or wing section will start to lose lift and this is when an wing or wing section stalls. The airflow over a wing can also separate at high Mach numbers, which is caused by shock waves that cause high pressure gradients [13]. This research will only focus on stall in the subsonic airspeed region as this is where currently all commercial airliners are flying. Only the Concorde and the Tupolev Tu-144 used to fly passengers beyond Mach 1. [13]

Subsonic stall can be split in three different types, one of which is the trailing edge stall discussed above. The other types of stall are thin airfoil stall and leading edge stall. The effects of the stall are vastly different and can be observed in Figure 2-2. Leading edge stall is characterized by a very sudden loss of lift, which can be seen in Figure 2-2b, as the flow separation at the leading edge suddenly causes flow separation over the entire wing. This is different from trailing edge stall, where the flow separation builds up more slowly. Finally, for thin airfoils, a separation bubble forms from the leading edge, which increases with increasing angle of attack [14]. A thin airfoil stalls when the separation bubble bursts.

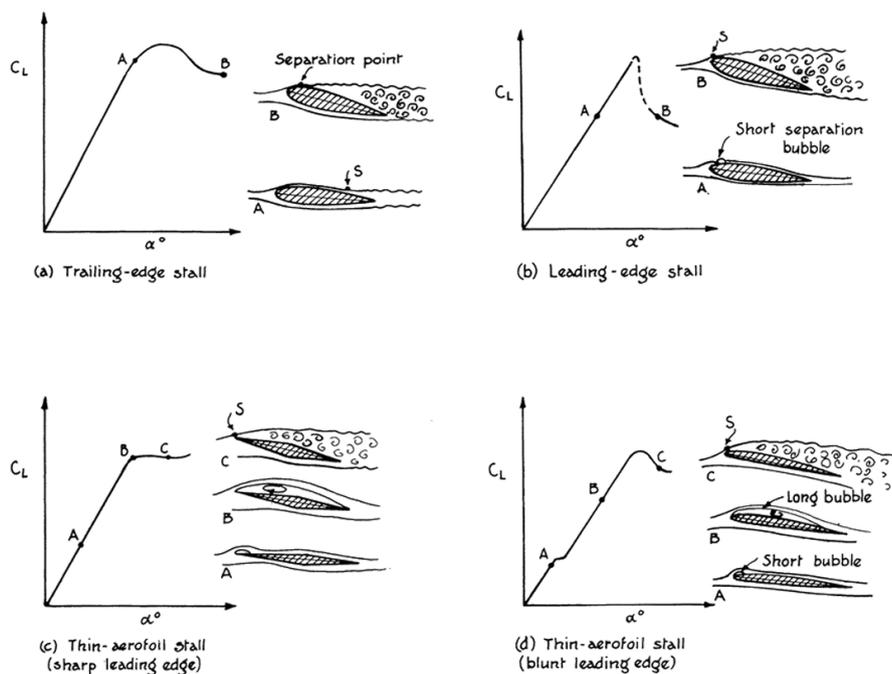


Figure 2-2: Three stall types [15]

Despite the fact that aerodynamic stall is highly nonlinear, for a given aircraft, it is only related to the angle of attack [16]. This is often a misconception, because many believe that stall is caused by a lack of speed, but this only because at a too low speed, the angle of attack is insufficient to provide enough lift. This makes the angle of attack an important parameter in the analysis of stall and it is used by ICAO [17] in the definition of stall: "An aerodynamic loss of lift caused by exceeding the critical angle of attack".

Because of the non-linearity of stall, no two stall scenarios are the same. However, despite the lack of repeatability, aerodynamic stall does have two very distinctive effects; a sudden

decrease in lift and an increase in drag [18]. These effects are the main basis of stall models and an important factor in stall modeling.

2-1-2 Stall Buffet

A stall buffet is the vibration of the aircraft due to the turbulent flow and its accompanied pressure fluctuations over the wings [19]. This is one of the most prominent warnings for an impending stall. The buffet starts before the stall, which can serve as an indication to pilots that they are nearing the critical angle of attack. It is therefore crucial to also model the buffet when modeling stalls.

2-2 Stall Models

Stall models are created to capture the aircraft behavior. These stall models can be used for analysis, simulation, and in simulators to support stall training. Stall models used in flight simulators are subject to several rules and regulations, which will be described first. After this, several stall modeling techniques are briefly described, followed by an extensive description of the TU Delft stall model and stall buffet model.

2-2-1 Stall Model Requirements

The International Civil Aviation Organization (ICAO) updated the regulations on upset prevention and recovery training for pilots. These regulations became active in 2019 and were implemented by the Federal Aviation Administration (FAA) and European Union Aviation Safety Agency (EASA). The regulations call for better training on upsets, which are generally defined by ICAO as "an undesired airplane state characterized by unintentional divergences from parameters normally experienced during operations" [17]. Stall is one of these upsets. Next to the mandatory training for pilots, the regulations on the stall model were also updated. The regulations of EASA [8] call for the following aspects that should be simulated in the stall model:

1. Degradation of the static/dynamic lateral-directional stability
2. Degradation in control response (pitch, roll, and yaw)
3. Uncommanded roll acceleration or roll-off requiring significant control deflection to counter
4. Apparent randomness or non-repeatability
5. Changes in pitch stability
6. Stall hysteresis
7. Mach effects
8. Stall buffet

9. Angle of attack rate effects

In order to certify a flight simulation training device (FSTD), a so-called statement of compliance (SOC) should be obtained, which confirms that "for each upset scenario, the recovery manoeuvre can be performed such that the FSTD does not exceed the FSTD training envelope, or when the envelope is exceeded, that the FSTD is within the realms of confidence in the simulation accuracy" [8]. This SOC is obtained by evaluation of a subject-matter expert, this is a pilot knowledgeable of the airplane's stall characteristics. The SOC should address the following aspects of the aerodynamic model: [8]

1. Source data and modeling methods used to create the aerodynamic model
2. Validity range of the aerodynamic model in terms of the angle of attack and side slip
3. How the model characteristics described above are addressed

The resulting SOC is used to certify an FSTD for training pilots. With the certified simulators, pilots can perform their stall training in several different scenario's mandated by the regulations. These scenario's include for instance stall at wings level, stall in turning flight of at least 25 degrees bank angle, and high altitude stall [20]. This requires that the stall model used in simulators is valid across all these conditions. This requires proper stall modeling, which will be discussed next.

2-2-2 Stall Modeling

In the past, there have been several attempts to model the stall dynamics of commercial aircraft. Fischenberg [21] was among the first to set up a stall model for the nonlinear dynamics beyond the critical angle of attack. As this model serves as a basis for the TU Delft stall model, this will be discussed separately in subsection 2-2-3.

Morelli, Cunningham, and Hill [22] propose a new model structure to represent the aircraft dynamics at higher angles of attack. They suggest to add specific coefficients that capture the dynamics beyond a certain angle of attack, and couple these to spline terms. A spline term, e.g. $(\alpha - 10)_+^2$, means that this term becomes "active" when its value is greater than zero, so in this case, when the angle of attack is larger than 10° , as can be seen in Equation (2-1).

$$(\alpha - 10)_+^2 = \begin{cases} 0, & \text{if } (\alpha - 10) \leq 0 \\ (\alpha - 10)^2, & \text{otherwise} \end{cases} \quad (2-1)$$

Morelli, Cunningham, and Hill [22] propose to use 8 of those spline terms to capture behavior beyond 8 different angles of attack (e.g. $(\alpha - 12.4)_+^0$, $(\alpha - 13)_+^1$, $(\alpha - 19)_+^2$, etc.).

Next to this, Ghazi et al. [23] uses a multi-layer perceptrons network to use the measured parameters to obtain the lift, drag, and moment coefficients. This is a type of neural network that they use to find a fit for the C_{Ls} , C_{Ds} , and C_{Ms} coefficient, based on an input layer that consists of measurable parameters such as the angle of attack, elevator deflection, and Mach number. The neural network is optimized for the given problem, by use of a genetic

algorithm. However, it still is more or less a "black-box" in which the dynamics are captured. This means that, although it produces the right flight trajectory, it is hard to say anything about the internal dynamics and how they are effected by different conditions.

Within the TU Delft, a task force was set up to develop a high fidelity aircraft stall model to be used in FSTDs. The first results were published by Van Horssen, De Visser, and Pool [24], who created the first model based on flight test data from TU Delft's research aircraft PH-LAB, a Cessna Citation II aircraft. Further efforts by Van Ingen, De Visser, and Pool [25] expanded the model by using an orthogonal model structure selection method and estimated the model parameters for that model.

2-2-3 TU Delft Stall Model

The stall model used the TU Delft is based on Kirchhoff's theory of flow separation and is used by [9, 10, 19, 24, 25]. This was first introduced by Fischenberg in 1995 [21]. He defined the X parameter as the point along the chord where flow separation took place. $X = 1$ means that the flow separation point is at the trailing edge, meaning no separation at all. Evidently, $X = 0$ means that the separation point is at the leading edge, meaning that the flow is fully separated. X is dependent on many factors, which Fischenberg described in his work [21]. He used Kirchhoff's theory to capture the circulation and boundary layer effects into a hysteresis factor, which resulted in Equation (2-2). Furthermore, he used a combination of Kirchhoff's theory and the Wagner or Theodorsen function to capture the unsteady aerodynamics and make the flow separation point time-dependent, which lead to the ordinary differential equation (ODE) as shown in Equation (2-3).

$$X(\alpha, \dot{\alpha}) = X_0(\alpha - \tau_2 \dot{\alpha}) \quad (2-2)$$

$$\tau_1 \frac{dX}{dt} + X = X_0(\alpha - \tau_2 \dot{\alpha}) \quad (2-3)$$

The X_0 parameter is the steady flow separation point and can be determined from a nonlinear function of the angle of attack. By using Equation (2-2), this parameter can be determined by using static wind tunnel data. However, when using flight data for identification of the parameter, Fischenberg [21] suggests to use an alternative approximation based on Equation (2-4). When combining all these equations together, the final equation for stall modeling can be constructed in Equation (2-5).

$$X_0 = \frac{1}{2} \{1 - \tanh(a_1 \cdot (\alpha - \alpha^*))\} \quad (2-4)$$

$$\tau_1 \frac{dX}{dt} + X = \frac{1}{2} \{1 - \tanh(a_1 \cdot (\alpha - \tau_2 \dot{\alpha} - \alpha^*))\} \quad (2-5)$$

The obtained X parameter is now dependent on time, as both α and $\dot{\alpha}$ are time-dependent. The X parameter can now be estimated from flight test data to determine the lift coefficient of the airplane by using Equation (2-6).

$$C_L(\alpha, X) = C_{L\alpha} \left(\frac{1 + \sqrt{X}}{2} \right)^2 \alpha \quad (2-6)$$

The different terms in Equation (2-5) represent the following, based on [21, 26, 27, 28]:

- a_1 [-] is a coefficient of the flow separation function. It is indicative of **the abruptness of the stall**, meaning that a higher value will lead to a more abrupt drop in lift and a sudden and quick flow separation. This can be seen in Figure 2-3, with an a_1 of 70. When the value for a_1 is lower, the flow separation is more gentle and as a result, the drop in lift is also less aggressive.
- α^* [rad] or [deg] is **the angle of attack for $X_0 = 0.5$** . A higher value of α^* means a delay in flow separation, leading to a higher critical angle of attack and a higher maximum lift coefficient. This effect can be seen in Figure 2-4.
- τ_1 [s] is a time constant. It captures **the transient effects of the stall**, as it is a factor that scales the time derivative of X . It influences the slope of the flow separation point and with that the lift coefficient. This can be seen in Figure 2-5, where the lift slope changes with different values for τ_1 when flow separation starts around 10° . It can also be seen that the flow separation point has a different slope with changing angle of attack.
- τ_2 [s] is also a time constant. It captures **the stall hysteresis effects** due to circulation and boundary layer effects that behave proportionally to the rate of change of angle of attack $\dot{\alpha}$. Hysteresis is the "a retardation (slow development) of an effect when the forces acting upon a body are changed" [29]. This can be seen in Figure 2-6, where a higher value of τ_2 results in an airflow that is longer attached with increasing angle of attack, thus delaying the flow separation with changing angle of attack.

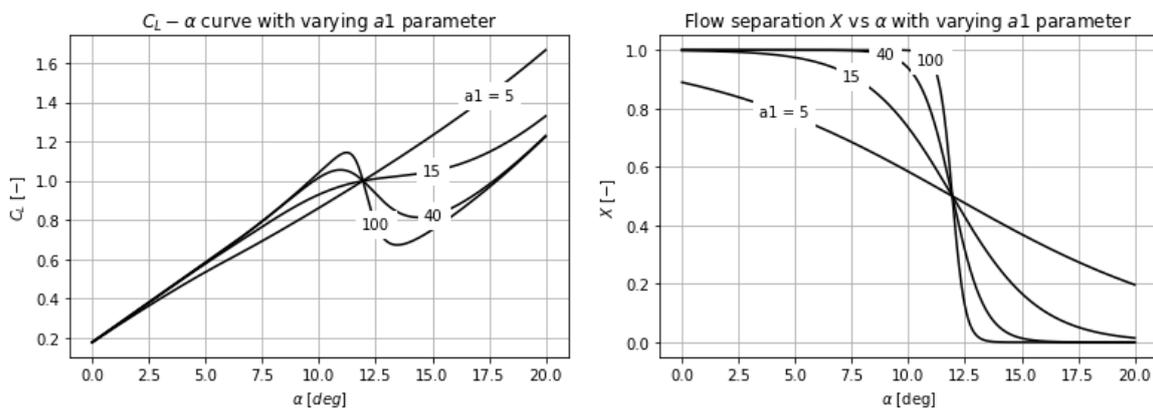


Figure 2-3: Effect of a_1 on the lift curve and flow separation point (adapted from [26])

Figures 2-3 to 2-6 have the same values for the Kirchhoff stall model parameters as found by [25], which can be seen in Table 2-1. Figures 2-3 and 2-4 are only analyzed in the static case, meaning that $\dot{\alpha} = 0$ and $\frac{dX}{dt} = 0$, so τ_1 and τ_2 have no influence and are set to zero in the

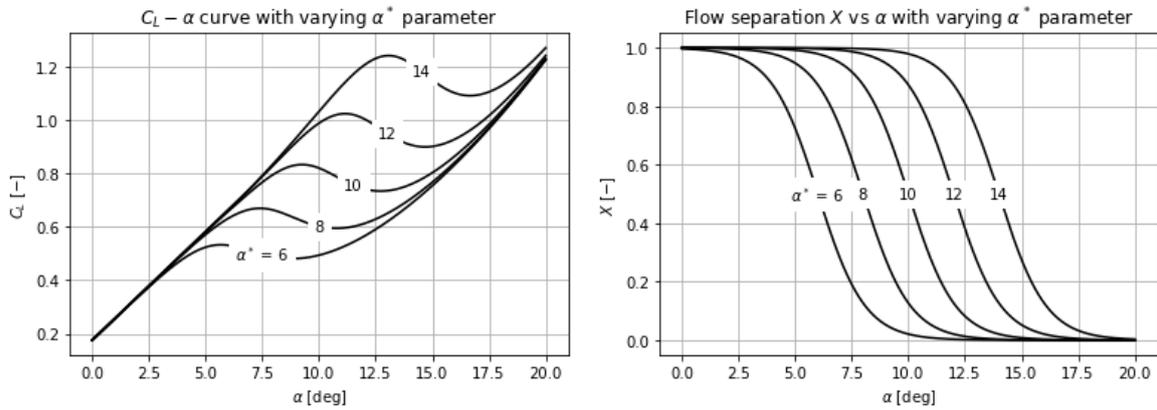


Figure 2-4: Effect of α^* on the lift curve and flow separation point (adapted from [26])

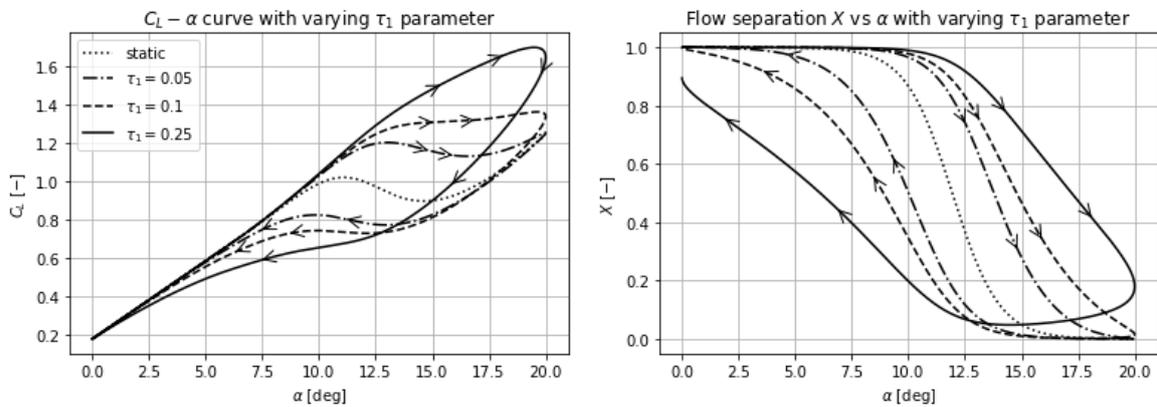


Figure 2-5: Effect of τ_1 on the lift curve and flow separation point (adapted from [26])

analysis. For Figure 2-5 and 2-6, a static and several dynamic cases are analyzed. For the dynamic case, Equation (2-7) is used for alpha, with $t = [0, 2]$ [sec].

$$\alpha(t) = 10^\circ + 10^\circ \cos\left(\frac{2\pi t}{2} + \pi\right) \quad (2-7)$$

The flow separation point X is used in the entire aircraft model. This obtained from Van Ingen, De Visser, and Pool [25], with the longitudinal model being:

$$C_L = C_{L_0} + C_{L_\alpha} \left(\frac{1 + \sqrt{X}}{2}\right)^2 \alpha + C_{L_{\alpha^2}} (\alpha - 6)_+^2 \quad (2-8)$$

$$C_D = C_{D_0} + C_{D_\alpha} \alpha + C_{D_{\delta_e}} \delta_e + C_{D_X} (1 - X) + C_{D_{C_T}} C_T \quad (2-9)$$

$$C_m = C_{m_0} + C_{m_\alpha} \alpha + C_{m_{X\delta_e}} \max\left(\frac{1}{2}, X\right) \delta_e + C_{m_{C_T}} C_T \quad (2-10)$$

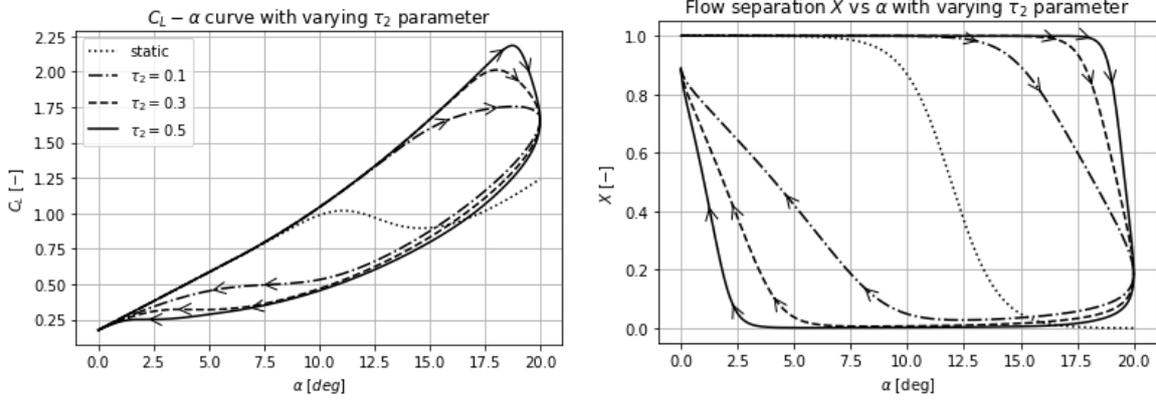


Figure 2-6: Effect of τ_2 on the lift curve and flow separation point (adapted from [26])

Table 2-1: Values used for creating Figures 2-3 to 2-6

| Parameter | Value | Unit |
|--------------------|---------|-------|
| C_{L_0} | 0.1758 | [-] |
| C_{L_α} | 4.6605 | [-] |
| $C_{L_{\alpha^2}}$ | 10.7753 | [-] |
| a_1 | 27.6711 | [-] |
| α^* | 0.2084 | [rad] |
| τ_1 | 0.2547 | [s] |
| τ_2 | 0.0176 | [s] |

The model structure for C_L is given in Equation (2-8) and makes use of a so-called spline term. This spline term becomes "active" when its value is greater than zero, so in this case, when the angle of attack is larger than 6° , as can be seen in Equation (2-11).

$$(\alpha - 6)_+^2 = \begin{cases} 0, & \text{if } (\alpha - 6) \leq 0 \\ (\alpha - 6)^2, & \text{otherwise} \end{cases} \quad (2-11)$$

The lateral model is given by the following equations:

$$C_Y = C_{Y_0} + C_{Y_\beta} \beta + C_{Y_p} \frac{pb}{2V} + C_{Y_r} \frac{rb}{2V} + C_{Y_{\delta_a}} \delta_a \quad (2-12)$$

$$C_l = C_{l_0} + C_{l_\beta} \beta + C_{l_p} \frac{pb}{2V} + C_{l_r} \frac{rb}{2V} + C_{l_{\delta_a}} \delta_a \quad (2-13)$$

$$C_n = C_{n_0} + C_{n_\beta} \beta + C_{n_r} \frac{rb}{2V} + C_{n_{\delta_r}} \delta_r \quad (2-14)$$

As can be noted from Equations (2-12) - (2-14), the flow separation point X is not a part of the lateral model. This was proposed in the work of Van Ingen, De Visser, and Pool [25] as their work suggested that this was not necessary. Currently, further research at the TU

Table 2-2: Proposed stall model parameters found by [25]

| Name | Value | Unit | Name | Value | Unit | Name | Value | Unit |
|--------------------|---------|-------|--------------------|---------|------|----------------------|---------|------|
| a_1 | 27.6711 | [-] | $C_{D(1-X)}$ | 0.0732 | [-] | C_{l_r} | 0.1412 | [-] |
| α^* | 0.2084 | [rad] | $C_{D_{C_T}}$ | 0.3788 | [-] | $C_{l_{\delta_a}}$ | -0.0853 | [-] |
| τ_1 | 0.2547 | [s] | C_{Y_0} | 0.0032 | [-] | C_{m_0} | 0.0183 | [-] |
| τ_2 | 0.0176 | [s] | C_{Y_β} | -0.5222 | [-] | C_{m_α} | -0.5683 | [-] |
| C_{L_0} | 0.1758 | [-] | C_{Y_p} | -0.5000 | [-] | $C_{m_{\delta_e X}}$ | -1.0230 | [-] |
| C_{L_α} | 4.6605 | [-] | C_{Y_r} | 0.8971 | [-] | $C_{m_{C_T}}$ | 0.1443 | [-] |
| $C_{L_{\alpha^2}}$ | 10.7753 | [-] | $C_{Y_{\delta_a}}$ | -0.2932 | [-] | C_{n_0} | 0.0013 | [-] |
| C_{D_0} | 0.0046 | [-] | C_{l_0} | -0.0017 | [-] | C_{n_β} | 0.0804 | [-] |
| C_{D_α} | 0.2372 | [-] | C_{l_β} | -0.0454 | [-] | C_{n_r} | -0.0496 | [-] |
| $C_{D_{\delta_e}}$ | -0.1857 | [-] | C_{l_p} | -0.1340 | [-] | $C_{n_{\delta_r}}$ | 0.0492 | [-] |

Delft is investigating the lateral stall model into more depth. At this moment, the model as proposed is deemed sufficient as this work will be focusing on symmetric stalls. The values of the coefficients from the proposed model can be seen in Table 2-2.

Recent, unpublished work [11] has focused on expanding the lateral stall model to better model stalls in a turn, both in coordinated and uncoordinated turns. However, this work is for now not considered as this research only focuses on quasi-steady symmetric flight.

Smets, De Visser, and Pool [9] investigated the sensitivity of the stall model parameters and found that a difference in α^* and a_1 had the most impact on the behavior of the aircraft during stall. Hence, these parameters are determined to be most influential in terms of their effects on the aircraft states. However, α^* is not a dynamic stall parameter which can be determined from wind tunnel data [21]. Hence, the parameter that has the most significant influence on the several aircraft states is a_1 . Furthermore, Smets, De Visser, and Pool [9] found that τ_1 variations have some influence and τ_2 variations have hardly any influence. This can be explained by the fact that τ_2 only has effect in the deep-stall regime which is hardly attained with the simulations done by Smets, De Visser, and Pool [9].

2-2-4 TU Delft stall buffet model

Van Horssen, De Visser, and Pool [24] proposed the vertical stall buffet model that is currently used by the TU Delft. It makes use of a white noise signal that is filtered through a second order filter, that can be seen in Equation (2-15) with the filter shown in Equation (2-16). A schematic overview of this can be seen in Figure 2-7, where $X_{\text{thres}} = 0.89$. The intensity of the buffet is scaled with $(1 - X)$, as a higher X means that the flow is more attached. The stall buffet model being the vertical buffet model, means that it only adds acceleration in the vertical direction. Van Horssen, De Visser, and Pool [24] also proposed a lateral stall buffet model. It has the same structure as the vertical stall buffet model, only with different parameters for the second order filters. These settings can be seen in Table 2-4.

$$S_{yy} = |H(j\omega)|^2 S_{uu} \quad (2-15)$$

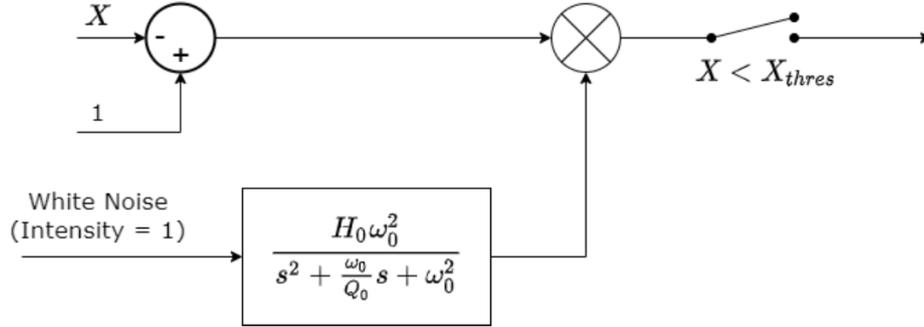


Figure 2-7: Schematic overview of the vertical stall buffet model as proposed in [24]

Table 2-3: Stall buffet model parameters for the vertical direction

| Parameter | Value | Unit |
|------------|-------|---------|
| H_0 | 0.05 | [-] |
| ω_0 | 75.92 | [rad/s] |
| Q_0 | 8.28 | [-] |

$$H(j\omega) = \frac{H_0 \omega_0^2}{(j\omega)^2 + \frac{\omega_0}{Q_0} j\omega + \omega_0^2} \quad (2-16)$$

2-3 Conclusions

When certifying a FSTD for stall training, a subject matter expert has to confirm that "the FSTD is within the realms of confidence in the simulation accuracy" [8] when the training envelope of the FSTD is exceeded. It is peculiar that there is no numeric required accuracy of the model that is used in FSTD beyond for instance the critical angle of attack. As a result, pilots all over the world are training in simulators that are certified based on the feelings of an expert instead of validated tolerances. Finding these tolerances for the stall model will be the objective of this research.

The stall model that will be used is developed by the TU Delft in previous research [24, 25] and is based on Kirchhoff's theory of flow separation as first introduced by Fischenberg [21]. This model uses a flow separation point X which can be calculated by solving the ODE as presented in Equation (2-5). The parameters shown in this equation can be determined from flight data, as was done by Van Ingen, De Visser, and Pool [25] for the TU Delft stall model

Table 2-4: Stall buffet model parameters for the lateral direction

| Parameter | Value | Unit | Parameter | Value | Unit |
|------------|-------|---------|------------|-------|---------|
| H_1 | 0.020 | [-] | H_2 | 0.01 | [-] |
| ω_1 | 36.43 | [rad/s] | ω_2 | 64.71 | [rad/s] |
| Q_1 | 4.19 | [-] | Q_2 | 11.99 | [-] |

that is currently in use. The X parameter as determined from the ODE can be used to update the aerodynamic model, as is shown in Figures (2-8) to (2-10).

The ODE has multiple parameters that are determined from flight tests. These parameters are a_1 , α^* , τ_1 , and τ_2 , and describe various stall characteristics. For instance, the a_1 parameter describes the abruptness of the stall. Smets, De Visser, and Pool [9] found that a_1 had the most influence on the different aircraft states. As a result, it is interesting to know the sensitivity of pilots with respect to changes in this parameter. In their analysis, Smets, De Visser, and Pool [9] already found that this parameter had the lowest threshold for a passive experiment, hence this parameter is most interesting to determine if the active thresholds differ from the passive thresholds.

Furthermore, the stall buffet as designed by Van Horssen, De Visser, and Pool [24] will also be added to the simulation. The stall buffet intensity is based on the value of X as determined by the ODE. This buffet model, together with the stall model will be implemented in the simulator in order to find the tolerances or accuracy needed. In order to say something about the required accuracy, one must first know the sensitivity of the human. This becomes a question of finding the human differential thresholds, which will be discussed in the next chapter.

Human Differential Thresholds

In order to say something about the sensitivity of humans to changes in system dynamics, one can look at the differential thresholds or more specifically, the Just Noticeable Difference (JND) threshold. This captures when one is not able to notice any difference between two scenarios. To measure this, there are several techniques which are discussed in section 3-2. As this experiment has human participants, there means that there are biases introduced by the participants. This is to be avoided as much as possible, and several techniques for this are discussed in section 3-3. During this thesis, the passive thresholds found in earlier research will be compared to active thresholds. The research that has been done in this will be discussed in section 3-4 after which the most important findings will be summarized in the conclusions.

3-1 Just Noticeable Difference Thresholds

Previous research at the TU Delft set out to find the perception thresholds of variations in stall dynamics [9]. They set up an experiment in which pilots would passively observe what happened and would determine based on what they saw and felt if there was a difference between two separate runs. They did this for the a_1 and τ_1 parameters of the TU Delft stall model discussed in subsection 2-2-3. They determined the threshold at which pilots would just notice a difference between the separate runs.

This is called the JND, which is a general phenomenon that can be applied in many ways. A rather straightforward example is weight. If one is given a bag that weighs 5 kilograms and one that weighs 4.99 kilograms, it will most likely be impossible for a human to notice the difference. However, if one bag weighs 5 kilograms and the other 2.5 kilograms, the difference will be easily noticeable. This means that somewhere between 2.5 and 0.01 kilograms will be the threshold where human will no longer feel the difference. This is the JND. Another example that reflects more on the current world affairs, is that JNDs are used in economics for determining for instance how much a price can rise without customers noticing a difference

[30], or by how much a package size can be reduced such that customers do not notice that they are paying the same amount for less of the product [31].

Although the JND examples seem rather straight forward, the measured JND for, for example the 5 kilogram weight, is not necessarily the same JND for a weight of 10 kilograms. However, there exists a relation between the JNDs of both weights, which is called Weber's Law [32]. Weber's law is described by Equation (3-1), where I is the reference property, also called the stimulus, ΔI_{jnd} is the JND, and c is a constant.

$$\frac{\Delta I_{\text{jnd}}}{I} = c \quad (3-1)$$

This allows for translating the JND of e.g. the 5 kilogram weight to the 10kg weight, by determining the value of the constant (also called the Weber fraction), and then applying this constant to the new reference weight of 10 kilograms.

The thing about these JNDs is that they are not absolute. The human sensitivity differs from moment to moment and is influenced by many factors [33]. Furthermore, it differs from person to person. As a result, the JND must be measured several times and with different people, which makes it a statistical concept [33]. There are several ways of measuring JND thresholds, which will be described next.

3-2 Measurement strategies

As Stevens says: "When we undertake to measure a JND, we in fact measure the confusions of the subject as he tries to cope with small stimulus differences." [33] Measuring JNDs can be done in several fashions, which will be described below. First, the method of constant stimuli is described, followed by a discussion on the normal staircase procedure, as these are the most basic methods of determining a threshold. After this, more advanced staircase procedures are described in subsection 3-2-3. This will be followed by an explanation of the Parameter Estimation by Sequential Testing method, after which the Maximum likelihood procedures are discussed in subsection 3-2-5.

3-2-1 Method of constant stimuli

The method of constant stimuli places several test conditions along the stimuli spectrum and keeps them fixed for each experiment. Consequently, it is a non-adaptive procedure [34]. Usually, there are 5 to 9 different stimuli levels chosen for the experiment. These stimuli should be placed between a stimulus level which is almost always detected and a stimulus level which is almost never detected [35].

The obtained results for each stimulus level will give the frequency describing how often a stimulus level is detected, which results in a discrete frequency distribution across all the different stimuli [34]. These frequencies will be used to obtain probabilities, which can be used to create a psychometric function. An example of a psychometric function can be seen in Figure 3-1. The psychometric function shows the relation between the level of stimulus and

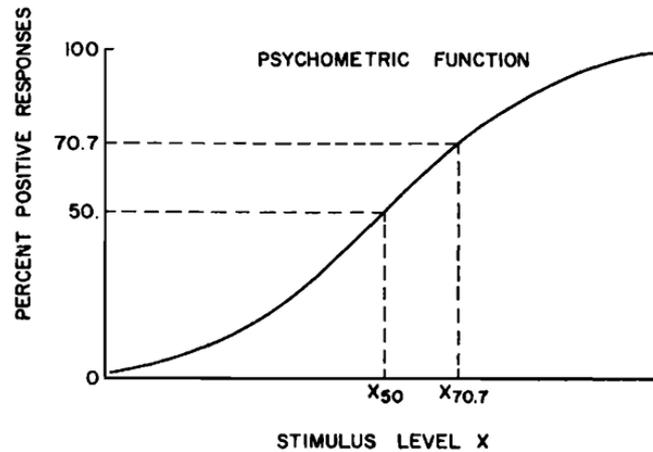


Figure 3-1: Example of a Psychometric Function [36]

its effect on the perception of participants by showing the increased proportion of positive answers, meaning a positive detection of the stimulus.

For the psychometric curve, often a Gaussian distribution is assumed [35], although a Weibull or logistic distribution is also possible [37]. Consequently, the psychometric curve can be fitted to the data based on the equation describing the Gaussian distribution, as can be seen in Equation (3-2) [9]. In this equation, σ and μ can be obtained from fitting the distribution to the data. The 50% threshold can be found by setting $P(\varphi) = 0.5$, which yields the value for φ , which is the stimulus level of the test parameter.

$$P(\varphi) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\varphi} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx \quad (3-2)$$

3-2-2 Staircase Procedure

In a staircase procedure, 4 different conditions should be determined before the start of the experiment [38]:

1. *Where to start* - this is an important step, because if the start of the experiment is far from the stimulus, the experiment requires many trials to determine the threshold
2. *Step-size* - If the step-size is too large, the answers 'Yes' and 'No' will alternate and no definitive conclusion about the threshold can be made. Consequently, if the step-size is too small, the experiment will take too long and the answers of the participant will not give as much information as when the step-size were larger.
3. *Stopping criterion* - Ideally, and in most cases also true, the different levels of stimuli change rapidly until an asymptotic level is reached. The results stay there as long as the conditions do not change. An idealization of this can be seen in Figure 3-2a. A more realistic result is shown in Figure 3-2b. In this case, more data points will mean a more reliable threshold value.

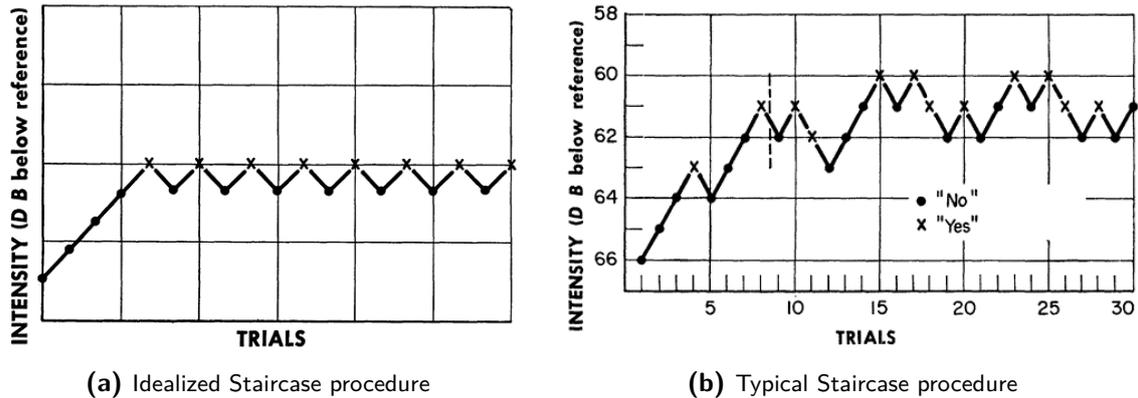


Figure 3-2: Different results of a staircase procedure [38]

4. *When to modify the step size* - Some experiments need a varying step-size, as described by Cornsweet [38], which is more commonly referred to as an adaptive staircase procedure.

The Staircase procedure yields the so-called 50% value, which is the stimulus value that is detected 50% of the time [38]. In other words, this is the stimulus value at which the chances of either a positive or negative answer are 50%. This phenomenon can be seen in Figure 3-1, where X_{50} is the stimulus value that is detected 50% of the time.

3-2-3 Advanced Staircase Procedures

The staircase procedure described above can be extended and changed slightly, with increasing performance and minor added complexity. A simple staircase procedure targets a 50% performance level. In order to increase performance, several different staircase procedures have been developed. One of these is the transformed staircase, which will be described first. After this, the weighted staircase will be described.

Transformed Staircase

A normal staircase approach takes a step up or down based on the answer of the previous level, sometimes described as a 1-down/1-up (1D/1U) procedure. The transformed staircase procedure is a more advanced approach of a normal staircase method. In this procedure, there are multiple consecutive steps at a consistent stimulus intensity level before taking a next step toward the threshold. During these successive steps at the same intensity level, a positive response must be given before advancing toward the threshold, while a negative response immediately results in a step away from the threshold.

An example of this is the 2-down/1-up (2D/1U) staircase procedure, which targets 70.7% performance level, which can be seen in Figure 3-1 as the $X_{70.7}$ point. This 2D/1U procedure means that 2 positive answers should be given before a next step down can be taken, whereas one negative answer reverses the steps back. There are more possibilities for this, such as a 3-down/1-up (3D/1U) staircase which targets 79.4% performance, or even a 4-down/1-up

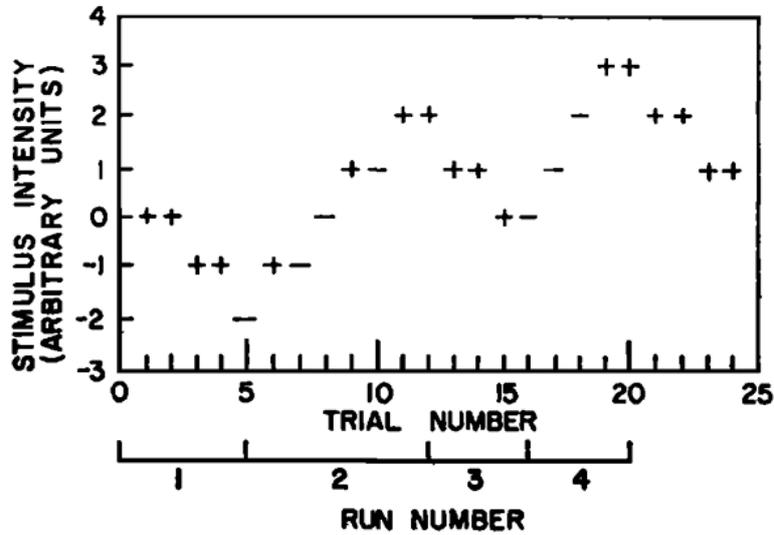


Figure 3-3: Example of a 2D/1U adaptive staircase procedure [36]

(4D/1U) staircase that has 84.1% performance. The principle can be seen in Figure 3-3, where a 2D/1U staircase procedure is shown. It can be seen that a step down is only taken when 3 positive answers have been given. However, when a negative answer has been given, a step up is taken.

Weighted Staircase

A weighted up/down method is based on the a normal staircase procedure with a 1-up/1-down steps. However, the step sizes for up and down are not the same. The ratio between them is determined by the targeted accuracy ψ_{target} and follows Equation (3-3). Reworking the equation gives the ratio between the step up and the step down, as can be seen in Equation (3-4) [39].

$$\psi_{target} = \frac{\Delta^+}{\Delta^+ + \Delta^-} \quad (3-3)$$

$$\frac{\Delta^-}{\Delta^+} = \frac{1 - \psi_{target}}{\psi_{target}} \quad (3-4)$$

For instance, if a targeted performance of 75% is desired, the ratio is determined as follows:

$$\frac{\Delta^-}{\Delta^+} = \frac{1 - 0.75}{0.75} = \frac{1}{3} \quad (3-5)$$

This means that for each negative answer, the stimulus level goes back up with 3 times the down step size. This can be seen in Figure 3-4. It is also possible to combine both the weighted staircase with the transformed staircase, which follows the formula given in Equation (3-6), where D is the number of consecutive correct responses before taking a step down [39].

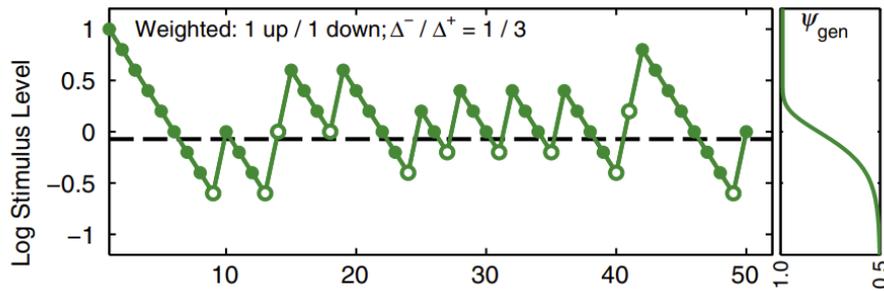


Figure 3-4: Weighted up/down staircase procedure with $\delta^-/\delta^+ = 1/3$ [39]

$$\psi_{\text{target}} = \left(\frac{\Delta^+}{\Delta^+ + \Delta^-} \right)^{\frac{1}{D}} \quad (3-6)$$

3-2-4 Parameter Estimation by Sequential Testing

The Parameter Estimation by Sequential Testing procedure, or PEST procedure in short, uses an adaptive approach towards threshold determination [40]. This adaptive approach uses the most ideal stimuli to converge towards thresholds quickly. There are a few things that need to be designed beforehand to set up a PEST procedure [41].

1. *When to change levels* - This can be set as a percentage. The experimenter needs to keep track of the total number of correct responses and the number of trials. With the help of a formula, the experimenter can determine when this percentage of correct responses is reached and determine if the current level should be run again or if the level can change.
2. *What level to try next* - Taylor and Creelman stated a few rules for this:
 - (a) "On every reversal of step direction, halve the step size" [41]
 - (b) "The second step in a given direction, if called for, is the same size as the first" [41]
 - (c) "On the fourth and subsequent steps in a given direction are each double their predecessor" [41]
 - (d) "Whether a third successive step in a given direction is the same as or double the second depends on the sequence of steps leading to the most recent reversal. If the step immediately preceding that reversal resulted from a doubling, then the third step is not doubled, while if the step leading to the most recent reversal was not the result of a doubling, then this third step is double the second." [41]
3. *When to stop* - The PEST is stopping at a predefined minimum step size. How small this step size is, determines the final precision of the estimated threshold but it hardly influences the efficiency of the procedure.

Finally, Taylor and Creelman [41] said that the threshold would simply be the final value of the PEST procedure. There were modifications proposed in later research, that used the entire data set to construct a psychometric function.

3-2-5 Maximum-Likelihood Procedures

A maximum-likelihood procedure utilizes the information of all measurements to determine the new settings for the next run. This allows for an even faster convergence. Close to the actual threshold, the procedure will behave more like a normal staircase procedure [42]. Before reaching close to the threshold, the procedure uses a maximum-likelihood estimation to determine the settings for the next run. Pentland [42] nicknamed this procedure the "Best PEST". The method is faster and converges to a more accurate threshold value. However, they warn that for more complex problems, this approach may become computationally heavy, especially when multiple variables are considered.

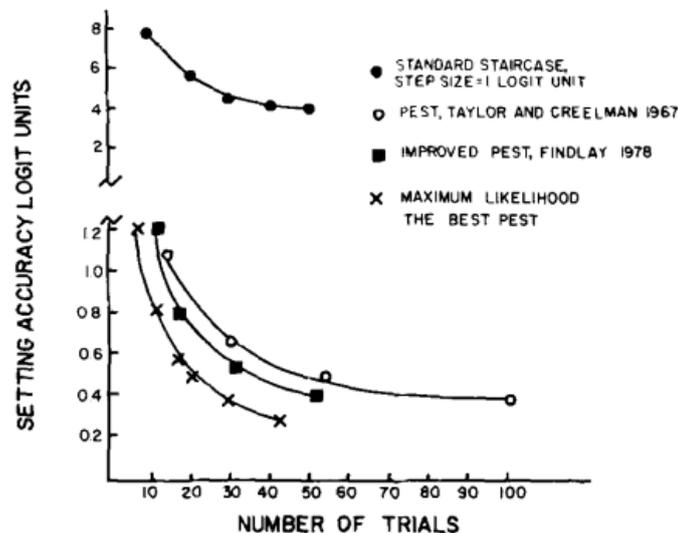


Figure 3-5: Comparison by Pentland [42] of several methods

Next to this, in order for the maximum-likelihood procedure to work, a psychometric function shape and slope value must be determined beforehand [40]. If the shape and slope are unknown, additional computations are needed before the maximum-likelihood procedures are used.

3-3 Biases

When working with human participants, it is almost inevitable that certain biases occur in the experiment. For instance, when asking if participants detected a stimulus, they might be tempted to answer "Yes" even though they did not detect a stimulus [9]. One way of preventing this, is by opting for a two-alternative forced-choice (2AFC) experiment. In a 2AFC experiment, participants are forced to make a decision. For example, when participants are asked to determine which of two light sources is brighter, they need to choose which of

the two options is the brightest, rather than to answer either "Yes" or "No" when asked if there is a difference.

With the 2AFC, the chance of getting a correct response is by default 50%. So, when participants are unsure which of the light sources is brighter, their answers will be 50% correct. When the difference between the two lights is clearly noticeable, their answers will be 100% correct. Somewhere between these two points, the JND threshold will lie. This can be seen in Figure 3-6, where a psychometric function is shown of a 2AFC method, as well as a 3AFC and 4AFC method. As the name suggests, 3AFC has three answer options, 4AFC has four options, which translate to a 33.333% and 25% correct responses by default, respectively.

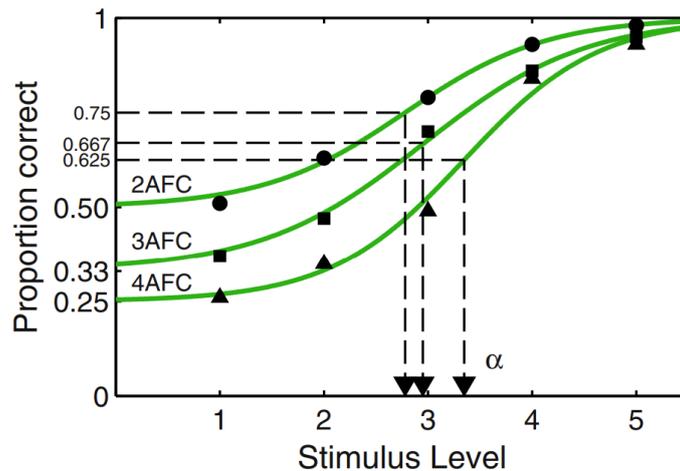


Figure 3-6: Example of a Psychometric Function with different measurement strategies (2AFC, 3AFC, and 4AFC) [39]

A 2AFC procedure is better than a yes/no question, as these "Yes/no tasks are particularly prone to the effects of bias" [39]. However, since certain procedures cannot be tested with 2AFC, another bias preventing technique was used by Smets, De Visser, and Pool [9]. They ran every third run with identical comparisons, which should ideally lead to a negative answer from the participants. When the participants were too often answering positively during these runs, they were most likely biased towards answering "Yes" and hence their results were excluded from the initial evaluation.

3-4 Passive vs Active Perception Thresholds

When revisiting Steven's quote "When we undertake to measure a JND, we in fact measure the confusions of the subject" [33], there is more to JNDs than only its measurement. The JND threshold is also influenced by the amount of confusion of the participant. This confusion can be influenced by many factors, one of which is the task a subject has to do during the experiment. In previous work by Smets, De Visser, and Pool [9], the JNDs of Kirchhoff stall parameters was measured with participants being passive in the simulator. A more realistic threshold can be obtained when the participants actually have to fly. By adding a control task, more confusion can be introduced. However, it also gives the participants an extra dimension to detect the differences.

Previous research has focused on the difference between active and passive thresholds, albeit for different thresholds. The work done by Hosman and Vaart [43] researched the perception thresholds of motion in pitch, while adding extra workload for the participants. The extra workload consisted of an auditory binary choice task and a roll control task. Although the participants had no influence on the pitch, the parameter which was measured, the experiment did show an increased threshold between 26% and 266%, depending on the different kind of added mental workload.

A similar result is found in the work by Pais et al. [44]. They also researched the difference between an active and passive JND threshold in pitch. They investigated the difference in threshold for pitch motion between an active control task and a passive observing role. In this research, the participants did influence their own pitch motion. In the research, it was hypothesized that the active threshold would be different from the passive threshold, however, without making any expectations on if it were to be higher or lower. The reasoning behind this was based on the fact that the participants would be controlling their own movement, which adds mental workload whilst giving the participants an extra cue to detect the threshold from. After the experiment, it was found that the absolute threshold of the participants was 60% higher in the active control task experiment.

3-5 Conclusions

A Just Noticeable Difference (JND) threshold is an indication of how much a stimulus can vary before a difference between the new and the original stimulus is noticeable. A JND threshold is not absolute, but rather a statistical concept [33]. This calls for multiple measurements with multiple participants. These measurements can be done in several different ways, from simple staircase procedures to rather complex procedures that optimize each next step of the procedure. This research will utilize a procedure called the Parameter Estimation by Sequential Testing (PEST), which is an adaptive staircase approach that can quickly converge towards a threshold.

Despite the efforts to create a procedure that carefully converges towards a threshold, the procedure as a whole may still be prone to biases from the participants. There are ways of asking a specific question, such that the answer is almost bias free: a two-alternative forced-choice (2AFC). This method is suitable for research in which a difference can be easily quantifiable, such as which bag is heavier, or which light is brighter. This could potentially work for this research too, as this research will focus on the a_1 stall parameter as discussed in section 2-3 which describes the stall abruptness. Therefore, asking for instance "*Which stall was more abrupt?*" creates a 2AFC method for this research.

However, initial testing of the experiment should indicate if asking for stall abruptness is even possible or if a question like "*Did you notice a difference?*" is the only way to go. As a result, this research might still use the method proposed by Smets, De Visser, and Pool [9]. This method will also involve the test runs where both settings are the same, which should give an indication on whether or not a participant may be biased to answer "*Yes*".

A different approach will be taken for the active control part of the experiment. For this experiment, the method of constant stimuli will be used. The reason for this is that active control can influence the results of a staircase procedure greatly. If one test run is flawed by

an error in the participant's control behavior, this might lead to results that are unexpected and can potentially damage the effectiveness of a staircase procedure. Therefore, to get an initial estimate of to what extent the passive thresholds are also applicable to active control, the method of constant stimuli is deemed to be a more robust method for this. The method of constant stimuli allows to construct a psychometric function, which can provide an insight in how the thresholds vary from passive to active, if they vary at all.

Finally, literature does provide an outlook on the possible results of this research. In previous work, a passive threshold was determined and compared to a threshold where an added mental workload or an active control tasks was added [43, 44]. Both researches reported an increase in threshold, of 27% to 266% and 60% respectively. Although the absolute threshold of pitch motion was determined, it suggests that the threshold of an active threshold of Kirchoff stall parameters might also be higher than those found in the work by Smets, De Visser, and Pool [9].

Stall in Flight Simulation Training Devices

Designing an experiment where participants fly themselves can be challenging, as their flying may cause certain biases in the outcomes. Such biases can greatly affect their ability to detect any differences between multiple runs, and hence, it is important to minimize the participant's own influence as much as possible. There have been several simulator experiments with stall, so there are several lessons that can be learned from them. This is described first, after which the research done at the TU Delft is discussed separately. Finally, the key takeaways are discussed in the conclusions of this chapter.

4-1 Earlier Stall Simulator research

There are a couple of research papers that elaborately describe the procedures followed during their stall experiments. These can serve as an inspiration for this research and are described below. First, the work by Schroeder et al. [45] is discussed, followed by a description of the work by Cunningham et al. [46]. After this the work by Grant, Moszczynski, and Schroeder [47] followed by brief summary of important aspects in other work found.

4-1-1 Work by Schroeder et al., (2014)

Schroeder et al. [45] evaluated several different stall models in a Level D 737-800 simulator. They evaluated 4 different stall models in different flight conditions. They had pilots flying the simulator themselves, for which they designed a method to ensure that the entry into the stall is as much the same for each run as possible. The pilots would establish a 1 kt/sec deceleration until the critical angle of attack. When the critical angle of attack was reached, a tone would sound in the cockpit and only then were the pilots allowed to recover from the stall. When the pilots were no longer descending, another tone would sound. They emphasized and

Table 4-1: Abbreviated Stall Recovery Template [45]

| | | |
|---|------------------------------------|---|
| 1 | Autopilot and auto throttle..... | Disconnect |
| 2 | a) Nose down pitch control..... | Apply until stall warning is eliminated |
| | b) Nose down pitch trim..... | As Needed |
| 3 | Bank..... | Wings Level |
| 4 | Thrust..... | As Needed |
| 5 | Speed brakes/Spoilers..... | Retract |
| 6 | Return to the desired flight path. | |

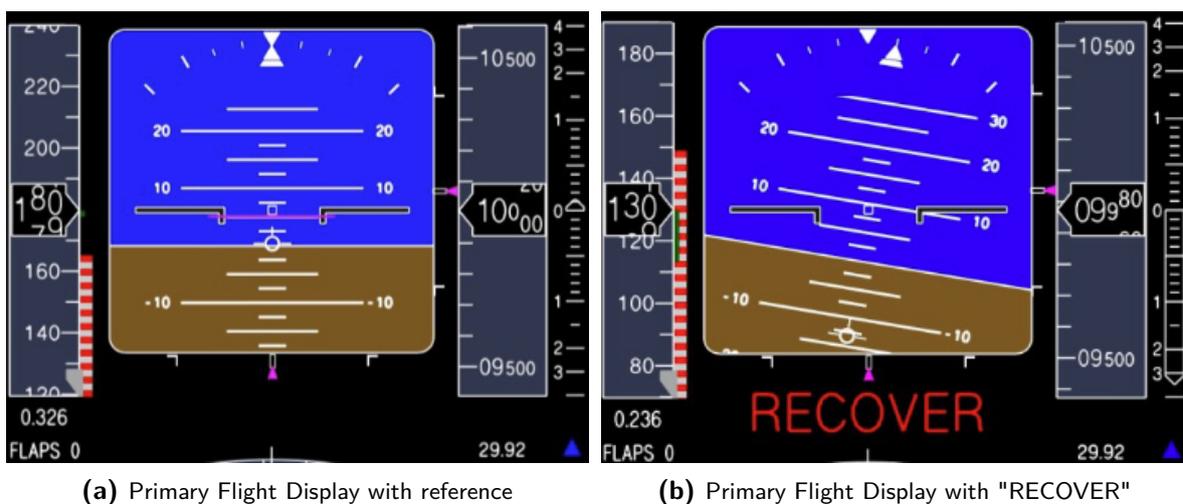
used the the standard FAA stall recovery procedures for the recovery phase, which can be seen in Table 4-1.

Furthermore, they disabled the stick shaker in the approach to stall phase, and would activate it again when the critical angle of attack was reached. This was done in order to minimize the possibility of negative training. This prevents that pilots practice with ignoring a stall warning.

4-1-2 Work by Cunningham et al., (2019)

Cunningham et al. [46] also conducted a stall model evaluation experiment, in which they assessed different flight dynamics model options and saw which were most important. They explicitly mentioned to the pilots that their stall recovery performance was not evaluated.

Cunningham et al. [46] used a different method to guide the pilots toward the stall angle of attack. They showed a pitch reference bar on the primary flight display (PFD), as can be seen in Figure 4-1a, that the pilots needed to follow. This reference bar would go from 5 to 10 degrees pitch attitude over a period of 13 seconds, after which the bar would disappear and the message "RECOVER" would be shown, indicating that they needed to start their recovery procedure.

**Figure 4-1:** Primary Flight Display used by Cunningham et al. [46]

4-1-3 Work by Grant, Moszczynski, and Schroeder, (2018)

Grant, Moszczynski, and Schroeder [47] investigated several different model fidelity effects on full stall recovery training. They also had an active flying experiment and took inspiration from the work by Schroeder et al. [45] by using a 1 kt/s deceleration into the stall. The model was based on a T-tailed turbo-prop aircraft and employed a stick pusher, which is commonly seen in T-tailed aircraft to prevent entering into a unrecoverable pitch up moment [48], also known as deep stall [49]. A deep stall is a strong form of stall which for a T-tail aircraft also greatly affects the performance of the elevator, meaning that stall recovery actions become ineffective resulting in the aircraft staying in stall [49].

Something that stands out from the procedures used by Grant, Moszczynski, and Schroeder [47] is that the pilots were told to steer the aircraft beyond the stick shaker. This is different from the procedures used by Schroeder et al. [45], who only activated the stick shaker when the pilots were to start their recovery procedure. They did this to prevent negative training.

Grant, Moszczynski, and Schroeder [47] also shed a light on the familiarization task. They mention that the pilots were allowed to get used to the cockpit and simulator via a test flight. During this flight, they were allowed to fly as they liked and were encouraged to enter turns and change speeds. The flight lasted about 5 minutes.

4-1-4 Other work

Lombaerts et al. [50] also performed a stall experiment where they used the autopilot to enter the stall in order to assure consistency across all runs and pilots. They used light turbulence in all their scenarios and started the flight in a trimmed state. Furthermore, they gathered data on the participants, such as total amount of flight hours, age, and which aircraft type they were currently certified for.

Stepanyan et al. [51] used another approach to let the pilots enter stall while flying themselves. They mimicked the Air Algérie Flight 5017 crash in Mali in 2014, where the autothrottle failed and set the throttle to its minimum before disengaging [52]. During the Air Algeria Flight, the autopilot was on altitude hold mode, and with the aircraft slowing down, the autopilot increased pitch to keep the altitude, eventually resulting in a stall. Stepanyan et al. [51] also set the throttle to minimum and instructed the participants to keep the altitude, which mimicked the autopilot on hold mode.

Finally, Schuet et al. [53] also designed an autopilot that would bring the aircraft into a stall. After the desired stall state was reached, the autopilot and autothrottle would disconnect and a RECOVER message would appear on the PFD, similar to what is shown in Figure 4-1b, and alarm would sound. They tested guidance systems to assist the participants during the stall recovery.

4-2 Earlier Stall Simulator Research at TU Delft

Previous research at the TU Delft has led up to this current research. This will be described below, first describing the work by Smets, De Visser, and Pool [9] followed by a short description of the work by Imbrechts, De Visser, and Pool [10].

4-2-1 Work by Smets, De Visser, and Pool, (2019)

Smets, De Visser, and Pool [9] tested the JND thresholds for key parameters of the Kirchhoff stall model mentioned before. Their experiment only focused on the thresholds for the τ_1 and a_1 parameters, as these were found to be most influential on the outcome of the simulation when changed. They investigated the JND thresholds through an adaptive staircase procedure, with the participants as passive observers. For this, staircase procedure, the direction of the steps was reversed after a different answer, e.g. reversed when "no" was answered after one or more consecutive "yes" answers. After a reversal, the step size was halved. This was only done after the second reversal, to prevent penalizing participants for an early mistake, which allowed for a relatively quick convergence. If the answer of a participant was four times the same in a row, the step size would be doubled for the next run. They called the procedure the 1-up/1-down adaptive staircase procedure, which yields a 50% correctness threshold. However, the correct name for this procedure is found to be a PEST procedure and follows the rules described in subsection 3-2-4.

Each trial consisted of two stall runs, one with the baseline model and the other one based on the previous answer. If the baseline model was used for the first or the second stall was done in random order. After two stall runs, they asked the question: *Did you notice a difference?*, after which the participant was to answer either "yes" or "no". This procedure can be seen in Figure 4-2.

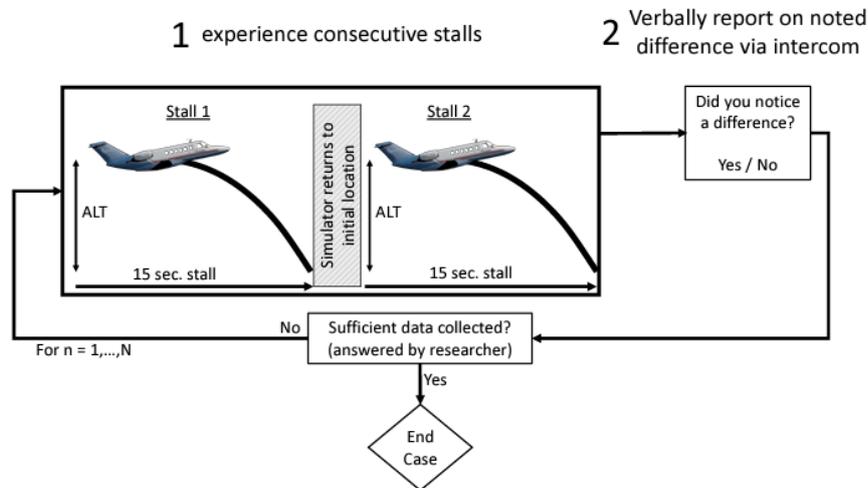


Figure 4-2: Experiment procedures of [9]

A "yes" or "no" question comes with drawbacks, as there is, for instance, not a correct and incorrect answer [37]. The participant simply expresses their subjective feeling. Because there is no correct answer, this procedure may lead to certain biases, called the response bias [37]. This can be because some people are, for instance, more inclined to answer yes or no. The threshold that is obtained by this procedure has a special name: the *point of subjective equality* [37].

In order to check the consistency of the participants, every third run of the experiment by Smets, De Visser, and Pool [9] had two times the baseline model, after which the participants should ideally answer that they did not feel a difference. If they often mentioned that they

felt a difference during this "null measurement" run, they were most likely biased towards answering yes and as a result, their data was not considered in the research. An example of this can be seen in Figure 4-3, where the data with the null measurements is shown in Figure 4-3a. The data without the null measurements can be seen in Figure 4-3b, where the reversals, the halving of the step size, and the final threshold can be seen. The stopping criterion was either a step size that was 1/32th of the first step, or when 30 trials (including the "null measurements") were performed.

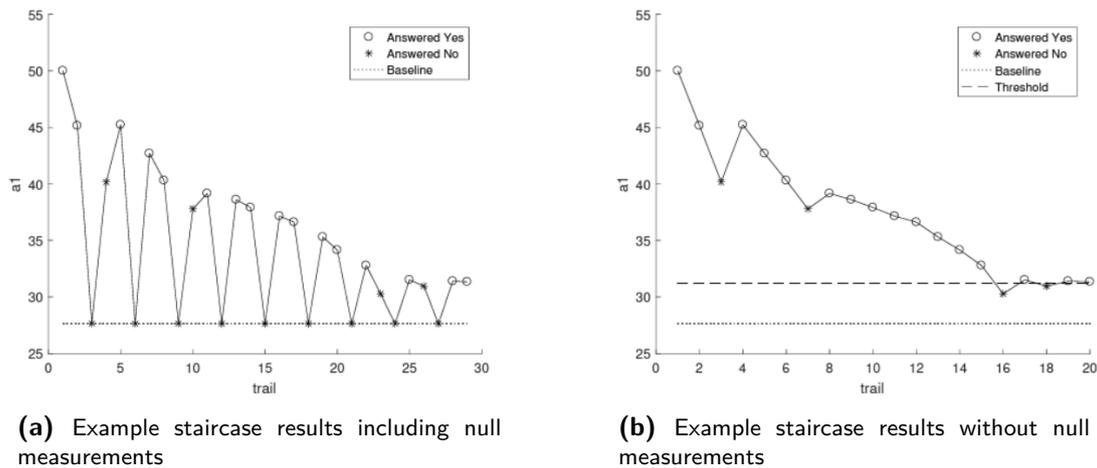


Figure 4-3: Example of the results obtained by Smets, De Visser, and Pool [9]

They conducted the experiment in the SIMONA research simulator at the Faculty of Aerospace Engineering of the TU Delft, which has a six degree-of-freedom hydraulic hexapod motion system. Their participants were passive observers, meaning that they did not fly themselves during the simulation. An autopilot recovered from the stall and the participants were asked if they felt a difference in the recovery. In order to prevent the participants from remembering certain values at the end of the simulation, the numeric indications of altitude and airspeed were removed from the PFD.

4-2-2 Work by Imbrechts, De Visser, and Pool, (2022)

Imbrechts, De Visser, and Pool [10] did a similar experiment as described in subsection 4-2-1 and investigated the JND thresholds for the stall buffet model parameters. They added a small, random parameter variation at every updated parameter setting in order to prevent participants from following identical paths along the stimuli levels. Furthermore, they used the same procedures as Smets, De Visser, and Pool [9].

4-3 Conclusions

For this research, a combination of the interesting aspects of the experiments described above will be used. This research will let the participants fly the approach to stall themselves. The pitch reference bar described in subsection 4-1-2 will be implemented to ensure that the initial approach to stall is similar for each experiment, and when entering the stall, the

"RECOVER" message will also appear on the primary flight display. This will be combined with the sound as used by Schroeder et al. [45] to make sure that the recovery starts at approximately the same moment during each test run. Finally, the participants will also be reminded and instructed to follow the FAA standard stall recovery procedure described in subsection 4-1-1, as this is the same procedure as from the EASA [54].

Furthermore, for the passive part of the experiment, the procedures followed by Smets, De Visser, and Pool [9] will be taken as a baseline and further improved where necessary. Moreover, the JND threshold found by Smets, De Visser, and Pool [9] will be used to set the different stimuli levels for the method of constant stimuli to be used in the active experiment.

Next to this, special attention will be paid to the familiarization task before the experiment, as the participants will have to get used to the controls and set up. Whereas Grant, Moszczynski, and Schroeder [47] only used a 5 minute familiarization flight, Smets, De Visser, and Pool [9] gave the participants about 15 minutes to familiarize themselves with the simulator, scenarios, and detecting differences between two test runs. The training scenario for this experiment will have to be designed in such a way to prevent biases, from either a over training or from a lack of familiarity with the simulator and controls.

Sensitivity Analysis

Before setting up the experiment and finalizing all details, it is good to know the limits of the simulation. These limits will influence the simulator settings and thus, it is important to get an estimate of where these limits lie. These can be found by performing several offline simulations. Next to this, through these simulations, it is possible to get a first estimate of the different control behavior of the participants and how this might influence the outcome of the experiment. For this sensitivity analysis, first the method will be discussed in section 5-1, after which the results will be given. Finally, the conclusions for this research will be presented.

5-1 Method

5-1-1 Set up

This sensitivity analysis is set up to perform a simulation of the Cessna Citation II of the TU Delft. During this simulation, the aircraft is kept wings-level, resulting in symmetric flight. This means that there are only 3 degrees-of-freedom (DOF) to analyze.

Several simulations will be performed to search for the limits of the simulation. This is done by changing a key parameter of the stall recovery auto pilot, in order to simulate different possible control behaviors of the participants. This is important to know, as the goal of the active experiment is to find the thresholds of changes in stall parameters and not to find the changes that the participants induced themselves.

As a result, the participants ideally behave like the stall autopilot that was used by Smets, De Visser, and Pool [9], i.e. a similar control strategy for each scenario such that only changes in stall parameters are felt and these are not influenced by a different control behavior. In order to see what the result of differences in control behavior could be, the stall autopilot as used by Smets, De Visser, and Pool [9] is taken and subtle changes to key parameters are introduced.

The key parameters identified from the stall autopilot are:

- The threshold $\alpha_{\text{threshold}}$ at which the recovery procedure is initiated
- The reference angle θ_{ref} during the stall recovery phase
- The controller gain P_{θ} of the reference angle θ_{ref}
- The threshold on \dot{h}_e when full thrust is applied
- The threshold on V_{tas} when the reference angle switches back to 10°

Changing these parameters can give an insight in how different pilots might react to the stall condition and how they recover. For instance, when a pilot is very quick with the controls, this can be simulated by increasing the controller P_{θ} as shown above. Similarly, when a nose-down input is applied, different pilots might aim for different references, which can be simulated by changing the reference angle θ_{ref} . Furthermore, pilots can have a different view on when the danger of a secondary stall is not present anymore and that it is therefore safe to return to the intended flight path. This can be simulated by changing the threshold on V_{tas} when the reference angle switches back to 10° .

The parameters as identified above are changed with a certain percentage with respect to the baseline. For α_{thres} , this is simulated with -15% - $+15\%$ with respect to the baseline value. A variation of α_{thres} with $\pm 15\%$ gives a range of $13.6^\circ - 18.4^\circ$. Larger variations would lead to an aircraft that is hardly stalling in case of a $< 13.5^\circ$ threshold variation or an unrealistic recovery scenario in case of a recovery threshold that is higher than 18.5° .

To keep an equal comparison as much as possible, the same percentage-wise variation was selected for \dot{h}_e . For V_{TAS} , this approach is not workable, as a 15% reduction of $86 [m/s]$ will result in a velocity of $73.1 [m/s]$. It is found that the stall speed for this simulation, at the specific height and corresponding flight conditions, $V_{\text{stall}} \approx 73 [m/s]$ (at $X = 0$). Therefore, applying a pitch up command at $V_{\text{TAS}} = 73.1 [m/s]$ to return to the original flight path, will result in a stall.

Therefore, a different approach is taken for V_{TAS} . The variation in V_{TAS} is taken with respect to the margin to stall, so with -15% - $+15\%$ of the difference between the baseline value and the stall speed of approximately $V_{\text{stall}} \approx 73 [m/s]$ at the specific height and corresponding flight conditions for this simulation. The difference between the baseline value and the stall speed is $86 - 73 = 13 [m/s]$. As a result, the variations for V_{TAS} are calculated as shown in Equation (5-1). All the parameters and their percentage-wise variations can be seen in Table 5-1.

$$V_{\text{test}} = 86 + x\% * (86 - 73) \quad (5-1)$$

Furthermore, for θ_{ref} and P_{θ} , a percentage-wise variation would not give meaningful results. Consequently, a range of seemingly realistic values are chosen for this analysis, which are shown in Table 5-2.

From the simulations, data on the trajectory and several important aircraft states is stored and analyzed. To analyze how different the aircraft behaved with the slightly changed parameters,

Table 5-1: Input parameters sensitivity analysis

| Parameter | -15% | -10% | -5% | Baseline | +5% | +10% | +15% |
|-------------------------------|-------|-------|-------|----------|-------|-------|-------|
| α_{thres} [rad] | 0.238 | 0.252 | 0.266 | 0.280 | 0.294 | 0.308 | 0.322 |
| h_e [m/s] | -15.3 | -16.2 | -17.1 | -18 | -18.9 | -19.8 | -20.7 |
| V_{TAS} [m/s] | 84.05 | 84.70 | 85.35 | 86.00 | 86.65 | 87.30 | 87.95 |

Table 5-2: Input parameters sensitivity analysis that would not benefit from percentage-wise step variation

| Parameter | | | | Baseline | | |
|-----------------------------|------|------|------|----------|-----|-----|
| θ_{ref} [deg] | -3.5 | -2.5 | -1.5 | -0.5 | 0.5 | 1.5 |
| P_{θ} [-] | | | 0.2 | 0.4 | 0.6 | 0.8 |

the variance accounted for (VAF) is used. The VAF compares two different signals and is a measure for how much similarity is in the data. A VAF of 100% means two completely the same signals. A lower VAF means less similarity. Theoretically, a VAF of $-\infty$ is possible. The VAF is calculated by the formula given in Equation (5-2). The VAF will be calculated of the states shown in Equation (5-3). Next to this, the changes with respect to the baseline model will be plotted and further investigated.

$$\text{VAF} = \left(1 - \frac{\sum_{i=1}^N (y(t_i) - \hat{y}(t_i))^2}{\sum_{i=1}^N y(t_i)^2} \right) \cdot 100\% \quad (5-2)$$

$$\Psi_{\text{VAF analysis}} = [q \ u \ v \ w \ \theta \ x \ y \ z \ X \ \delta_e \ \alpha] \quad (5-3)$$

The second part of the simulation will focus on seeking the limits of the simulation with respect to the Kirchhoff stall parameters. By simulating the most extreme conditions of the different simulated human control behaviors together with different settings of the Kirchhoff stall model, the JND experiments can be simulated to find the extremes in the simulation, which can be accounted for when setting up the simulator. These simulations will be done in similar fashion as the simulations mentioned before, however, with different Kirchhoff stall parameter settings. These settings can be seen in Table 5-3. It is based on the JND threshold found by Smets, De Visser, and Pool [9]. The different scenarios can be seen in Table 5-4.

Table 5-3: Input parameters of sensitivity analysis part 2, with the Baseline values from [9]

| Parameter | -2*JND | -1*JND | Baseline | +1*JND | +2*JND |
|-----------|---------|---------|----------|---------|---------|
| a_1 | 23.8802 | 25.7756 | 27.6711 | 31.9574 | 36.2436 |
| τ_1 | 0.1263 | 0.1905 | 0.2547 | 0.3351 | 0.4155 |

5-1-2 Simulation

For the simulation, the stall model created by [24, 25] is used. The model is created in a Matlab/Simulink environment in which the Cessna Citation II is simulated, the aircraft which is also the research aircraft of the Faculty of Aerospace Engineering at TU Delft (see

Table 5-4: Different scenarios sensitivity analysis part 2

| # | \mathbf{P} | θ_{ref} | \mathbf{a}_1 | τ_1 | # | \mathbf{P} | θ_{ref} | \mathbf{a}_1 | τ_1 |
|---|--------------|-----------------------|----------------|----------|----|--------------|-----------------------|----------------|----------|
| 1 | 0.2 | -3.5 | 27.6711 | Var | 7 | 0.4 | 1.5 | 27.6711 | Var |
| 2 | 0.2 | -3.5 | Var | 0.2547 | 8 | 0.4 | 1.5 | Var | 0.2547 |
| 3 | 0.2 | 1.5 | 27.6711 | Var | 9 | 0.8 | -3.5 | 27.6711 | Var |
| 4 | 0.2 | 1.5 | Var | 0.2547 | 10 | 0.8 | -3.5 | Var | 0.2547 |
| 5 | 0.4 | -3.5 | 27.6711 | Var | 11 | 0.8 | 1.5 | 27.6711 | Var |
| 6 | 0.4 | -3.5 | Var | 0.2547 | 11 | 0.8 | 1.5 | Var | 0.2547 |

Figure 5-1). The model has several options to simulate the flight, which includes trimming the aircraft for steady straight symmetric flight. These are the flight conditions that will be used at the start of the simulation, with the Citation II trimmed. The inputs and trimmed states are given in Table 5-5 and Table 5-6, respectively.

**Figure 5-1:** The Cessna Citation II research aircraft PH-LAB [55]

Furthermore, the autopilot controls are altered in the Simulink model and the "stall autopilot" as described by Smets, De Visser, and Pool [9] is implemented. This autopilot brings the aircraft into a stall with a deceleration of about 1 kt/s . When the angle of attack $\alpha \geq 0.28 \text{ rad}$, the recovery control loop of the autopilot will take over and bring the nose of the aircraft down. When the change in altitude $\dot{h}_e \leq -18 \text{ m/s}$, the autopilot will apply maximum thrust to the engines to bring the aircraft back to its original height. Furthermore, the roll attitude controller keeps the aircraft symmetric. The full scheme of the autopilot can be seen in Figure A-1.

5-2 Results

The sensitivity analysis consists of two parts, which will each be discussed separately. Firstly, the results of different autopilot settings, mimicking different pilot control behavior, will be shown, followed by the discussion on the sensitivity of the Kirchhoff stall model parameters to different pilot control behavior.

| Parameter | input |
|----------------|----------|
| δ_a | -0.031 |
| δ_e | -0.052 |
| δ_r | -0.060 |
| δ_{t_e} | 0.000 |
| δ_{t_a} | 0.000 |
| δ_{t_r} | 0.000 |
| δ_f | 0.000 |
| landgear | 0.000 |
| T_n no. 1 | 2254.163 |
| T_n no. 2 | 2254.163 |

Table 5-5: Inputs for trim

| Parameter | states | derivatives |
|-------------------|----------|------------------|
| p_{body} | 0.000 | $4.15193e - 12$ |
| q_{body} | 0.000 | $-8.77352e - 12$ |
| r_{body} | 0.000 | $-6.03531e - 12$ |
| V_{TAS} | 90.000 | $-2.83366e - 11$ |
| α | 0.123 | $-6.67075e - 12$ |
| β | 0.021 | $-3.17500e - 12$ |
| ϕ | 0.000 | $0.00000e + 00$ |
| θ | 0.123 | $0.00000e + 00$ |
| ψ | 0.000 | $0.00000e + 00$ |
| h_e | 5620.000 | $8.99807e + 01$ |
| x_e | 0.000 | $1.86542e + 00$ |
| y_e | 0.000 | $0.00000e + 00$ |

Table 5-6: Trim values for the simulation runs

5-2-1 Sensitivity of autopilot

The results for the various simulation runs can be seen in Tables 5-7, 5-8, and 5-9. The VAF results for changes in $\alpha_{\text{threshold}}$ suggest that there is quite a significant change in aircraft behavior with varying autopilot behavior. Especially the behavior of δ_e , which is the parameter that the pilots directly control, has a VAF of at minimum 52%. However, this result is mainly due to a time-shift in the signal due to a delayed start of the recovery, as can be seen in Figure A-2. The differences in aircraft states between the baseline and the results of the $\alpha_{\text{threshold}}$ variations can be observed in Figure 5-2. The actual states of the baseline and the variation can be seen in Figure A-2.

For comparison, two simulation results are added to the graphs, one with the a_1^+ threshold and the other with the a_1^- threshold as found by Smets, De Visser, and Pool [9]. These thresholds are represented by the two red lines, visible in for instance Figure 5-2. Between these red lines, the variations should go unnoticed as they indicate the boundary of the JNDs. Beyond these lines, the variations can be detectable. The most important parameters where the pilot variations are outside the JND thresholds, are the z-position of the aircraft and the pitch angle θ . This indicates that the simulation has a noticeable difference in attitude and altitude during the simulation, and that the final position in the simulation is also different. This can actually have significant influence on the results of the question "*Did you notice a difference?*". Smets, De Visser, and Pool [9] argue that one of the parameters their participants used to determine if there was a difference or not, was the outside visual of the simulator.

Very little differences can be observed in the other pilot induced variations. For instance, the results of the different h_e settings did not result in any changes in the behavior of the aircraft. For completeness, the results that were found can be seen in Table B-1 and Figures A-5 and A-10, where all parameters have a 100% VAF, meaning that they are all the same. The change in V_{TAS} did not give results that lie outside of the indicated JND thresholds, as can be seen in Figures A-6 and A-11 and Table 5-10. Neither the altitude z nor the pitch angle θ deviate significantly beyond the JND threshold.

Finally, a similar pattern can be found for the two parameters that did not vary percentage

wise. Both the reference angle of attack during the recovery θ_{ref} and the gain P_θ show changes in aircraft states that lie within the JND thresholds and should therefore not be detectable. Furthermore, a VAF of within 99% for all states can be observed for both signals in Tables 5-8 and 5-9, indicating high similarity between all variations. Consequently, these parameters are expected to not influence the results of the experiment too greatly.

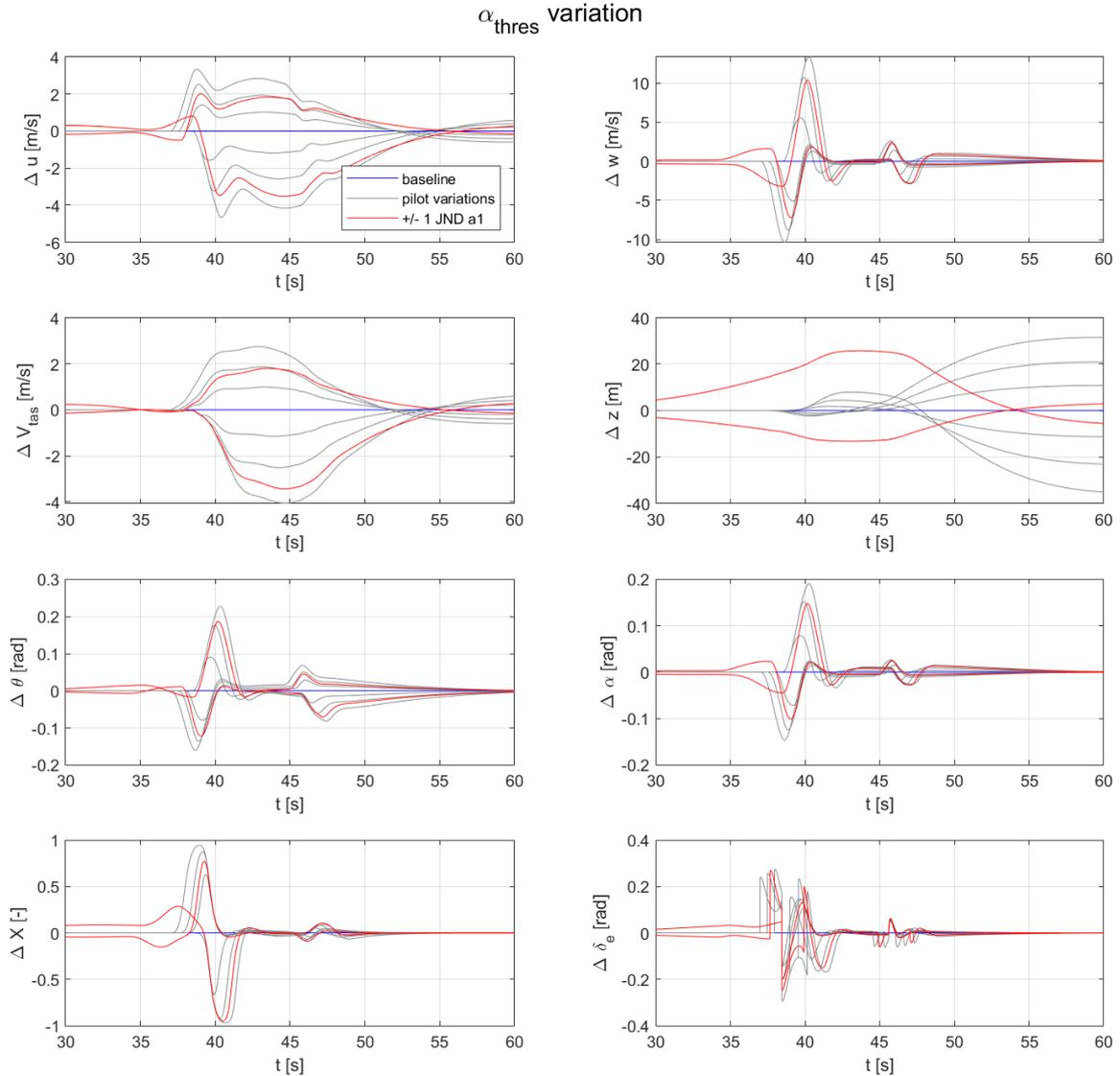


Figure 5-2: Difference for all states between baseline and variations in $\alpha_{\text{threshold}}$

5-2-2 Effects of different human control strategies

For the experiment, it is important that the simulator settings remain the same. However, before the motion system settings can be set the same, it is important to know that the different control strategies of the participants do not cause the motion system to reach its

Table 5-7: VAF with respect to $\alpha_{\text{threshold}} = 0.28$ [rad]

| | $\alpha_{\text{threshold}}$ [rad] | | | | | |
|------------|-----------------------------------|--------|--------|--------|--------|--------|
| | -15% | -10% | -5% | +5% | +10% | +15% |
| q | 99.906 | 99.956 | 99.996 | 99.984 | 99.831 | 99.694 |
| u | 99.797 | 98.742 | 99.727 | 99.788 | 99.259 | 98.425 |
| v | 99.615 | 98.614 | 99.719 | 99.806 | 99.260 | 98.361 |
| w | 31.555 | 63.703 | 90.153 | 92.176 | 76.864 | 62.123 |
| θ | 78.861 | 90.037 | 97.515 | 98.107 | 94.224 | 89.919 |
| x | 99.996 | 99.999 | 100 | 100 | 99.999 | 99.998 |
| y | 99.694 | 99.831 | 99.981 | 99.996 | 99.956 | 99.906 |
| z | 96.694 | 98.548 | 99.663 | 99.703 | 98.893 | 97.512 |
| X | 58.802 | 76.021 | 92.217 | 92.932 | 79.497 | 64.443 |
| δ_e | 52.659 | 54.887 | 70.645 | 75.347 | 66.477 | 68.618 |
| α | 56.810 | 77.366 | 93.904 | 95.159 | 85.625 | 76.307 |
| Avg | 86.865 | 90.910 | 95.728 | 94.865 | 87.053 | 78.909 |

limits. This is what is simulated in this part of the sensitivity analysis. The results can be seen in section B-2.

The simulations for the different a_1 settings of up to ± 2 JND show that the simulations stay between the maximum and minimum a_1 value that Smets, De Visser, and Pool [9] have set for their experiment. As a result, the pilot induced variation is expected not to lead to the simulator reaching the limits of its motion system. When the experiment is prepared in the simulator, a separate Heave-Gouverneur analysis will be done to fully confirm this. This will then be compared to the previous analysis by Smets, De Visser, and Pool [9].

5-3 Conclusions

As discussed, different settings for $\alpha_{\text{threshold}}$ show that the behavior of the aircraft can differ significantly, showing behavior that lie outside of the JND threshold as found in previous research. This means that, when participants are actively flying, they can influence the results of the experiment significantly, which can lead to a difference being detected whereas this difference is influenced only by their own control behavior. As a result, it is desirable to restrict the freedom to start the recovery procedure at any angle of attack, and thus fix the entry into stall procedure up until the desired recovery point. This can result in a less biased result. This means that the start of the recovery procedure will become a control variable instead of leaving it up to the participant to initiate the recovery.

Table 5-8: VAF w.r.t. $\theta_{\text{ref_recovery}} = -0.5$ [deg]

| | $\theta_{\text{ref_recovery}}$ [deg] | | | | |
|------------|---------------------------------------|--------|--------|--------|--------|
| | -3.5 | -2.5 | -1.5 | 0.5 | 1.5 |
| q | 99.999 | 99.999 | 100 | 99.997 | 99.999 |
| u | 99.618 | 99.829 | 99.956 | 99.823 | 99.618 |
| v | 99.581 | 99.812 | 99.952 | 99.805 | 99.581 |
| w | 98.412 | 99.284 | 99.821 | 99.328 | 98.412 |
| θ | 99.031 | 99.569 | 99.892 | 99.587 | 99.031 |
| x | 99.999 | 99.999 | 100 | 99.999 | 99.999 |
| y | 99.997 | 99.999 | 100 | 100 | 99.997 |
| z | 99.817 | 99.919 | 99.980 | 99.922 | 99.817 |
| X | 99.758 | 99.887 | 99.970 | 99.859 | 99.758 |
| δ_e | 98.339 | 98.881 | 99.544 | 99.096 | 98.339 |
| α | 99.225 | 99.650 | 99.912 | 99.661 | 99.225 |
| Avg | 99.434 | 99.712 | 99.912 | 99.914 | 99.734 |

Table 5-9: VAF w.r.t. gain $P_\theta = 0.4$

| | Gain P_θ [-] | | |
|--|---------------------|--------|--------|
| | 0.2 | 0.6 | 0.8 |
| | 99.999 | 99.999 | 99.999 |
| | 99.841 | 99.932 | 99.808 |
| | 99.824 | 99.923 | 99.786 |
| | 99.227 | 99.553 | 98.552 |
| | 99.534 | 99.777 | 99.333 |
| | 99.999 | 100 | 99.999 |
| | 99.999 | 100 | 100 |
| | 99.911 | 99.964 | 99.902 |
| | 99.805 | 99.850 | 99.469 |
| | 99.105 | 99.390 | 98.423 |
| | 99.670 | 99.765 | 99.227 |
| | 99.714 | 99.832 | 99.500 |

Table 5-10: VAF w.r.t. $V_{\text{TAS}} = 86$ [m/s]

| | V_{TAS} | | | | | |
|----------------|------------------|-------|-------|--------|--------|--------|
| | -15% | -10% | -5% | +5% | +10% | +15% |
| q | 99.999 | 100 | 100 | 99.999 | 99.999 | 99.998 |
| u | 99.84 | 99.93 | 99.98 | 99.98 | 99.92 | 99.59 |
| v | 99.82 | 99.92 | 99.98 | 99.98 | 99.91 | 99.54 |
| w | 98.43 | 99.19 | 99.77 | 99.75 | 99.05 | 95.35 |
| θ | 99.49 | 99.76 | 99.93 | 99.93 | 99.71 | 98.49 |
| x | 100 | 100 | 100 | 100 | 100 | 100 |
| y | 100 | 100 | 100 | 100 | 100 | 100 |
| z | 99.94 | 99.97 | 99.99 | 99.99 | 99.97 | 99.87 |
| X | 99.71 | 99.87 | 99.97 | 99.97 | 99.89 | 99.80 |
| δ_e | 98.87 | 98.86 | 99.23 | 99.22 | 98.70 | 97.64 |
| α | 99.31 | 99.65 | 99.90 | 99.89 | 99.60 | 98.08 |
| Average | 99.58 | 99.74 | 99.89 | 99.88 | 99.70 | 98.94 |

Experiment Setup

After reviewing the available literature and performing a preliminary sensitivity analysis, the actual experiment proposal for this thesis can be set up. From the research question, two hypotheses follow. These hypotheses are used to determine the experiment variables, from which the experiment design itself follows.

6-1 Experiment Hypotheses

The research question to be answered in this thesis is:

To what extent are the just noticeable difference thresholds of the Kirchhoff stall model parameters measured during a passive experiment also representative for the thresholds of pilots flying actively in a flight simulation training device?

From this, there are two hypotheses which will be tested for this experiment.

1. The active flying JND thresholds of the Kirchhoff stall model parameters will be higher than the passive thresholds found by Smets, De Visser, and Pool [9].
2. The upper passive threshold for a_1 will be the same as found by Smets, De Visser, and Pool [9], at $a_1^+ = 0.1549 \pm 0.1394$

The first hypotheses is based on the previous research regarding the translation of passive thresholds to active thresholds, as described in section 3-4. Although the research described there was measuring an absolute threshold rather than a JND threshold, it is the only source of information found in which passive and active thresholds are compared. The previous research noted an increase in threshold when the participants were actively controlling their motion. As this is also the case in this experiment, the hypotheses is made that the active

thresholds will be higher than the passive thresholds found by Smets, De Visser, and Pool [9].

To be able make useful comment on this transfer of passive to active thresholds, the passive thresholds of the participants of this experiment will be measured too. This allows for a second hypotheses to be made about these thresholds. Finding these passive thresholds can confirm the previous research carried out by Smets, De Visser, and Pool [9], which directly translate to the second hypotheses stating that the passive thresholds that will be measured during this experiment will be the same as found before.

6-2 Experiment Variables

From the hypotheses, the experiment design can be made. What follows first from the hypotheses are the independent variables. Next to that, the dependent variables can be chosen. Finally, the control variable are set in order to prevent confounding factors as much as possible in this experiment.

6-2-1 Independent Variables

From the hypotheses, only one independent variable comes: the passive versus active comparison. This calls for two different experiments, one in which the passive threshold will be determined for the participants and a second one that measures the active threshold. The experiment that determines the passive threshold will use the procedures as outlined by Smets, De Visser, and Pool [9]. This experiment will be referred to as the 'Passive Experiment'.

For the experiment that determines the active threshold, most of the experiment will be the same as the passive experiment, with the exception of the role of the participant and the settings for a_1 . This will be further discussed in subsection 6-2-3. This experiment will be referred to as the 'Active Experiment'.

6-2-2 Dependent Variables

The dependent variables of this experiment are whether or not the participants felt a difference. For the passive experiment, this will simply be the answer to the question "*Did you notice a difference?*". For the active experiment, this question might result in biased answers. Because the participants can actually induce differences in the flight themselves, asking if they felt a difference is not sufficient.

For this, the participants will be asked: "*Did you feel a difference in the simulator settings?*". This might steer participants away from evaluating their own induced differences and focus on the underlying simulator (model) settings. Furthermore, this will hint the participants towards using their command of the aircraft to their benefit in detecting the differences.

However, reformulating this question might also introduce biases into this research. By asking for instance to detect differences in the simulator model, already the expected results of the experiment become clear to the participants. Therefore, it is chosen to ask for differences in *simulator settings*. This generalizes the source of the differences, which can mean anything

from motion system settings to the model settings which we are actually interested in. This also keeps the actual goal of this research a bit more hidden, which is desirable.

Furthermore, the available data from the simulator will be recorded. This will help to evaluate the test runs in hind-sight, to see whether the difference felt by the participants was actually a difference or a difference induced by themselves.

6-2-3 Control Variables

Finally, the control variables are set. These are what is kept constant throughout the experiment in order to prevent their influence confounding the experiment.

- **SIMONA research simulator:** both experiments will be hosted inside the SIMONA research simulator (or SIMONA in short). The inside and outside of SIMONA is shown in Figure 6-1
- **Left-hand seat:** The participants will all be placed in the left-hand side of the cockpit, as this is the side which has a control column instead of a side-stick, as can be seen in Figure 6-1b. The control column is more representative of the actual Cessna Citation II aircraft. This is different from the experiment one by Smets, De Visser, and Pool [9], where the participants were seated in the right-hand seat due to the set up of SIMONA at the time of the experiment. However, this change is assumed to have little influence on the results of the passive experiment.
- **Motion settings of SIMONA:** for both experiments, the motion settings will be kept constant, and will be the same as the experiment of Smets, De Visser, and Pool [9]. It is expected that the limits of the motion system will not be reached, even with participants actively controlling the aircraft, as was already shown in subsection 5-2-2.
- **Instrument panel:** The instrument panel will be kept the same in both experiments, meaning that it will show the same information to the participants in each scenario.
- **SIMONA cabin:** All other cabin settings in the SIMONA will be kept the same for both experiments.
- **Initial condition:** Each simulation will start at the same point in the simulation environment, with the same initial settings for wind, turbulence, airspeed, altitude, trim conditions, etc.
- **Entry into stall:** As mentioned before in subsection 5-2-1, the entry into the stall will be fixed for each run. For the passive experiments, this is already incorporated in the autopilot, however for the active experiments this requires a guide for the participants to follow. The guide suggested for this is given in subsection 4-1-2.
- a_1 **settings:** This variable will be controlled in different ways for the two experiments
 - **Passive Experiment:** During the passive experiment, the staircase procedure as proposed by Smets, De Visser, and Pool [9] will be used.

Table 6-1: Active experiment settings for a_1

| Condition | a_1 Setting | a_1 Value | # of repetitions |
|-----------|---------------|-------------|------------------|
| 1 | Baseline | 27.6711 | 6x |
| 2 | + 0.5 JND | 29.8142 | 4x |
| 3 | + 0.75 JND | 30.8858 | 4x |
| 4 | + 1.0 JND | 31.9574 | 4x |
| 5 | + 1.5 JND | 34.1005 | 4x |
| 6 | + 2.0 JND | 36.2436 | 4x |
| 7 | + 2.5 JND | 38.3869 | 4x |

**(a)** Outside of SIMONA**(b)** Inside of SIMONA**Figure 6-1:** SIMONA research simulator [9]

– **Active Experiment:** During the active experiment, a set of conditions will be tested. This allows to get more data around the threshold. The set of conditions can be seen in Table 6-1. These settings will be tested before the actual experiment, and refined where necessary. Similarly to the passive experiment, the baseline setting will also be part of the trials. This is done in order to check the consistency and potential bias of the participants. If a participant often detects a difference during two similar runs, this means that

- **Stall autopilot:** For the passive experiment, the autopilot settings will remain the same.
- **Stall buffet:** For every simulation run, the stall buffet settings will be the same.

6-3 Other Experiment Design Variables

Next to the normal design variables of any experiment, there are several other factors that should be chosen for this experiment. Firstly, the participants for this experiment will be chosen, followed by the condition sequence that these participants will follow. Then, the

experiment as the participants will experience is discussed, followed by a description of the background procedures.

6-3-1 Participants

The participants will be pilots who have at least an active multi-engine pilot license. Mostly, this will mean that the potential participants are active certified commercial airline pilots. Next to this group, the group of Cessna Citation II pilots will be invited to partake in this experiment as well. The reason for choosing multi-engine licensed pilots is that the participants need to be able to control a stalling aircraft which has two engines. For this reason, glider pilots are not considered to be suitable for the experiment, as their approach to stall and stall recovery procedures are different to the approach and recovery of pilots who have an engine at their disposal.

Furthermore, the reason for excluding pilots who have a private pilot license (PPL) with a single engine aircraft, is that the single engine is usually a piston engine, which has a different handling than the dual turbofan engines of the Cessna Citation II. Finally, commercial pilots are the one who most often fly stalls in simulators, making them the best test group for the goal of this research.

The Cessna Citation II pilots that are connected to the research aircraft the TU Delft jointly operates with the Netherlands Aerospace Center (NLR) will also be asked. This group consists of 6 pilots, who, ideally, all do this experiment. Furthermore, the amount of commercial pilots that are desired for this experiment is at least 14, which results in a test group of 20 pilots. As seen in the research by Smets, De Visser, and Pool [9], some pilot may not perform as consistent and are therefore biased. In their research, they excluded 4 out of 12 of their participants. As a result, their work was based on 8 participants. With this same logic, it can be expected that up to 7 participants may not provide unbiased results. Seeing that test groups of around 12 participants are generally still accepted in previous research [46, 56], a test group of 20 participants is deemed acceptable.

6-3-2 Condition Sequence

In this experiment, there are only two different test cases that each participant will do. In order to prevent correlation effects to shadow the results, half of the test group will first start with the passive experiment, followed by the active experiment. The other half of the participants will start with the active experiment first. Each participant will be given a unique participant number, which will be a sequential range. Every odd-numbered participant will start with the passive experiment first, every even-numbered participant will start with the active experiment first.

The Cessna Citation II pilots will be given a different number than the other participants, in order to also evenly distribute their results. As a result, half of the Cessna Citation II pilots will start with the passive experiment, and the other half will start with the active experiment, following the same logic as the other group. The logic for both groups is presented in Table 6-2.

For the passive experiment, the next test condition (i.e. the a_1 setting) is determined by the answer to the question *"Did you feel a difference in the simulator settings?"* of the previous

Table 6-2: Logic for experiment order for each participant, with N the number of other pilots

| Cessna Citation II pilots | Other pilots | Order of experiment |
|---------------------------|-----------------|---------------------|
| Pilots A, C, E, ... | Pilots $2N - 1$ | Passive, Active |
| Pilots B, D, F, ... | Pilots $2N$ | Active, Passive |

run. For the active experiment, there are several predetermined conditions, as shown in Table 6-1, with each of these conditions having several repetitions. For this, a Latin Square procedure is set up. Because there are 7 conditions, a Latin Square of 14 is set up in order to balance all conditions. This can be seen in Table 6-3. However, since there are 2 extra baseline conditions, another Latin Square is created to balance this too. This is shown in Table 6-4. Here, there are 4 Latin Square conditions that will be taken from Table 6-3 and two extra baseline conditions. All test conditions will be implemented in the software and based on a participant number, the right condition sequence will be selected.

Table 6-3: Latin Square design for active experiment conditions

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|----------|----------|----------|----------|----------|----------|----------|
| a | Baseline | +0.5JND | +0.75JND | +1.5JND | +1JND | +2.5JND | +2JND |
| b | +1JND | +2JND | +0.75JND | +2.5JND | Baseline | +1.5JND | +0.5JND |
| c | +1.5JND | +2.5JND | +0.5JND | +2JND | Baseline | +1JND | +0.75JND |
| d | Baseline | +0.75JND | +0.5JND | +1JND | +1.5JND | +2JND | +2.5JND |
| e | +2JND | +1JND | +2.5JND | +0.75JND | +1.5JND | Baseline | +0.5JND |
| f | +1.5JND | +0.5JND | +2.5JND | Baseline | +2JND | +0.75JND | +1JND |
| g | +0.75JND | Baseline | +1JND | +0.5JND | +2JND | +1.5JND | +2.5JND |
| h | +2JND | +2.5JND | +1JND | +1.5JND | +0.75JND | +0.5JND | Baseline |
| i | +0.5JND | +1.5JND | Baseline | +2.5JND | +0.75JND | +2JND | +1JND |
| j | +0.75JND | +1JND | Baseline | +2JND | +0.5JND | +2.5JND | +1.5JND |
| k | +2.5JND | +2JND | +1.5JND | +1JND | +0.5JND | +0.75JND | Baseline |
| l | +0.5JND | Baseline | +1.5JND | +0.75JND | +2.5JND | +1JND | +2JND |
| m | +1JND | +0.75JND | +2JND | Baseline | +2.5JND | +0.5JND | +1.5JND |
| n | +2.5JND | +1.5JND | +2JND | +0.5JND | +1JND | Baseline | +0.75JND |

Table 6-4: Latin Square design with extra baseline conditions taken into account

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|----------|----------|----------|----------|----------|----------|
| A | LS 1 | LS 2 | LS 4 | Baseline | Baseline | LS 3 |
| B | LS 2 | Baseline | LS 1 | LS 3 | LS 4 | Baseline |
| C | Baseline | LS 3 | LS 2 | Baseline | LS 1 | LS 4 |
| D | LS 3 | Baseline | Baseline | LS 4 | LS 2 | LS 1 |
| E | Baseline | LS 4 | LS 3 | LS 1 | Baseline | LS 2 |
| F | LS 4 | LS 1 | Baseline | LS 2 | LS 3 | Baseline |

6-3-3 Experience Participants

Before starting the experiment, the participants will be thoroughly briefed on the experiment, the experiment structure, their task, and the safety features of SIMONA. After this, they will sign a consent form, that is confirming their voluntary participation, their consent to process certain personal data in accordance with the GDPR, and that they received the safety briefing. The briefing of the experiment will be in written, to ensure that each participant receives the same briefing and the experimenter does not incidentally spoil anything of the experiment.

When all is clear to the participant, they will be brought to SIMONA and again the safety features will be shown. After this, the participant will take place in the left-hand side of the cockpit and fasten their seat belts. They will wear a noise-canceling headphone through which they can communicate with the experimenter. When the participant is ready, they will start with familiarization runs.

For the active experiment, the familiarization run will involve already controlling the aircraft and bringing the aircraft into stalls by following the reference bar. When a consistent recovery is achieved, they will have a couple of runs in which they can detect differences between the baseline model and a test run with a τ_1 setting that is above the JND threshold found by Smets, De Visser, and Pool [9]. This to prevent training too much on detecting a_1 differences already. There will also be a couple of comparisons with no differences.

For the passive experiment, the familiarization run will show a couple of runs with the autopilot controlling the aircraft, followed by a couple of similar test runs in which the participant can detect the differences with the autopilot on. In both scenario's, when the participant feels confident that they can start the experiment, the actual experiment will start.

The participants are told that both experiments will follow the logic shown in Figure 6-2. After one of the experiments is completed, the participant will have a break. Then, they will continue with the other experiment, which will have the training and experiment as discussed before.

6-3-4 Background Procedures during the Experiment

As can be seen in Figure 6-2, there are two subsequent stalls during each run. One of these will always be the baseline model, and the other will be selected a_1 condition, or also the baseline model when a consistency check is done. If the baseline model is run on the first or second run, will be chosen randomly by the software. The baseline model will be the model proposed by Van Ingen, De Visser, and Pool [25], which was discussed in subsection 2-2-3. The other run will have the same model, only with a different a_1 setting.

For the passive experiment, the background procedures selecting the next a_1 setting will be the same as done by Smets, De Visser, and Pool [9]. This means that each participant will start at an a_1 value of 50, with an initial step size of 5. With each positive answer to the question "*Did you feel a difference in the simulator settings?*", the next a_1 setting will be closer to the baseline value, as determined by the staircase procedure. The staircase procedure that is used is the PEST, as described in subsection 3-2-4.

The PEST procedure has the following rules, as outlined before:

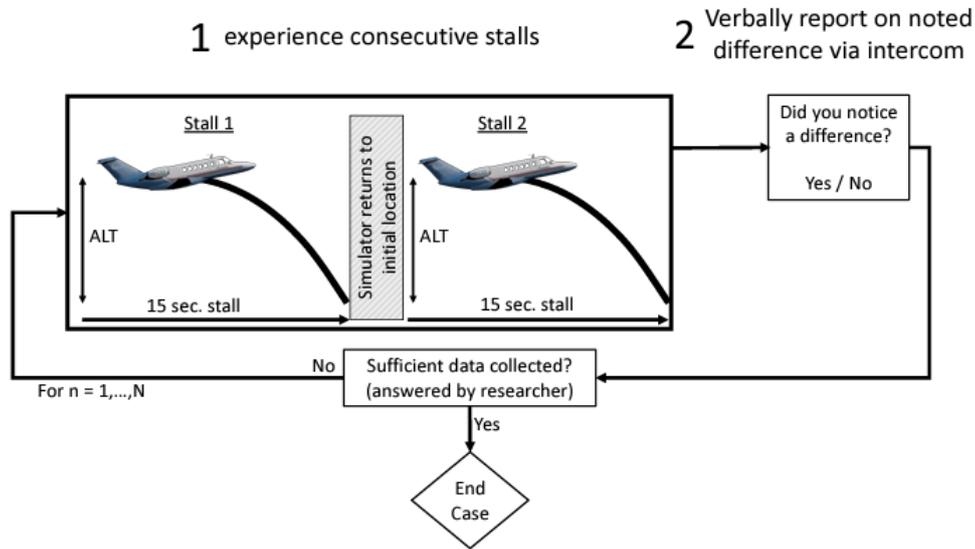


Figure 6-2: Experiment structure as designed by [9]

1. *When to change levels:* A 1-up/1-down procedure is used, meaning that each comparison changes the level based on the answer
2. *What level to try next:* The next level to try is based on rules 2a-d as shown in subsection 3-2-4. The only addition is that the first reversal will not half the step size, as this might give a slower convergence if a mistake is made in the initial phase of the experiment.
3. *When to stop:* This will be either when the step size has reached $1/32$ nd of the initial step size of 5 or when 30 trials has been reached.

This procedure with these rules will be implemented in the software, such that all steps will be automated based on the responses of the participants.

For the active experiment, the conditions as given in Table 6-1. Still, despite the procedure being set up front, the participants will not be informed on how many trials they will be taking. This will have them believe that they are following the procedure set in Figure 6-2 and that the trials will stop until sufficient data has been collected.

6-4 Final considerations

Now that the experiment is set up based on literature and preliminary simulations, the next task is to implement all this in SIMONA. However, the preliminary simulations and literature only provide limited insight in how the experiment will actually feel inside the SIMONA. As a result, a number of aspects of the experiment design are as of yet uncertain and will require testing before the decisions are final. The first of these decisions is to make it an 2AFC experiment, which will be discussed in subsection 6-4-1, followed by a discussion on the staircase procedure of the experiment. Finally, the thoughts experiment conditions of the active experiment are mentioned.

6-4-1 Two-alternative forced choice design

Because this experiment will only change a_1^+ , this experiment allows for a 2AFC design. A 2AFC design is different from a Yes/No-question as was explained in section 3-3, and forces the participants to make a choice between two options rather than ask if they felt a difference. Because a_1 represents the stall abruptness, the question to be asked can be *Which model was more abrupt?*. However, this question removes the bias that the participants may have in a Yes/No question which, in turn, removes the need to have the measurements that Smets, De Visser, and Pool [9] performed to check the consistency of the participants. So now, instead of dedicating 1/3 of the measurements to detecting biases, all measurements can contribute towards finding the threshold.

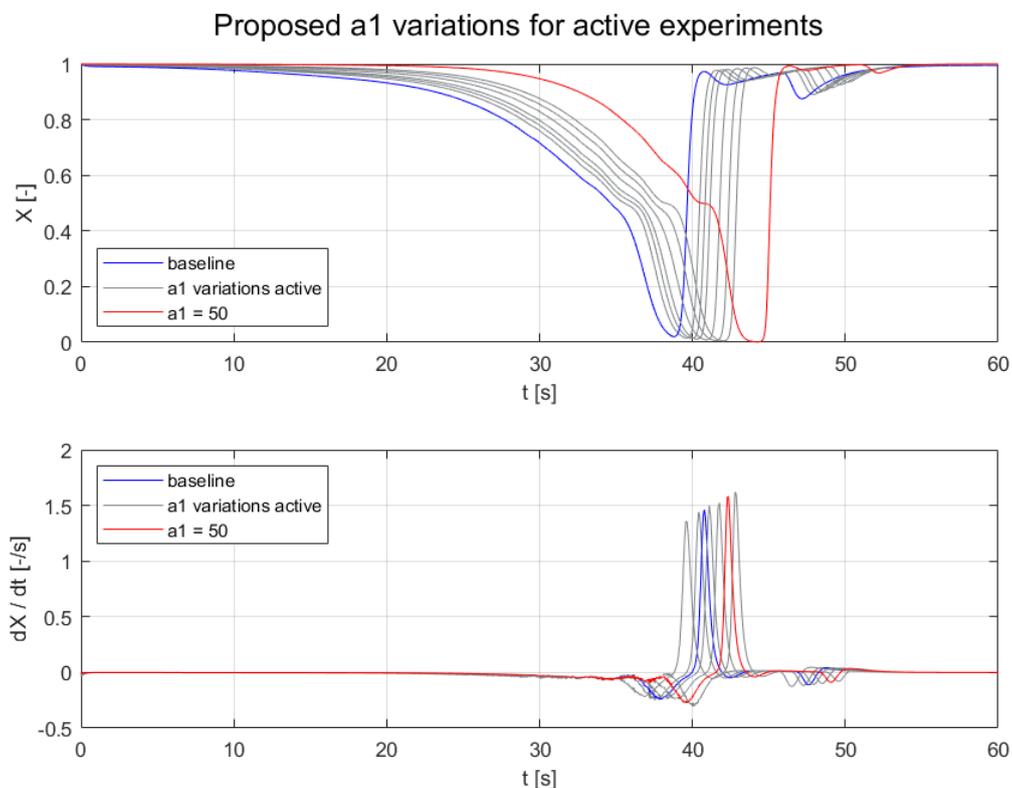


Figure 6-3: X and $\frac{dX}{dt}$ for the active experiment conditions as well as the maximum a_1 value for the passive experiment

Some preliminary simulations are done in order to see back in the data the results of this, which can be seen in Figure 6-3. Here, the change in abruptness becomes visible. The difference between the dX/dt graphs seems to be rather small. As a result, initial testing once SIMONA is set up must provide an answer to whether or not it is possible to transform the experiment to a 2AFC experiment and ask which simulation was more abrupt. For reference, the other states of this simulation are shown in Figure A-18.

6-4-2 Staircase for passive experiment

As highlighted in chapter 3, there are numerous ways of setting up the staircase procedure. The design that is currently chosen is the PEST procedure, as is highlighted in subsection 6-3-4. However, there is still a possibility to switch to a 2 Down/1 Up (2D/1U) design. In this design, a level will be repeated twice before moving closer to the threshold. And with a negative response, in either the first or second run of the condition, the condition will immediately move back a step, away from the baseline.

However, with a 2D/1U PEST procedure, the targeted correctness level changes. With a 1 Down/1 Up (1D/1U) design, the targeted correctness level is 50%, and with a 2D/1U design, the correctness level is 70.7%. This can be explained with a mathematical approach. If $p(T)$ is the probability of a correct response, the probability of two correct responses in a row is $[p(T)]^2$. The threshold is defined to be the point where the decision to decrease the stimulus is just as likely as the decision to increase it, which makes $[p(T)]^2 = 0.5$, which yields $p(T) = 0.707$ [37].

It would be beneficial to implement a 2D/1U procedure, as it can give more certainty in the data, especially close to the threshold. For the 1D/1U, if a participant is not sure whether or not they felt a difference and they opt to say "Yes", this means that during the next run, the stimulus goes down. However, in a 2D/1U procedure, the participant has two trials to make sure that they felt a difference before actually stepping down to the next level. This is also beneficial in a 2AFC design. If a participant guesses correctly in a 1D/1U procedure, the stimulus level goes down immediately, whereas in a 2D/1U procedure, the next trial at the same stimulus level serves as an extra layer of "protection" against correct guess responses.

There are also drawbacks for implementing a 2D/1U procedure. By testing each level twice, the experiment will take longer and might therefore influence the participants' concentration towards the end of the experiment. Furthermore, a difference in correctness level means that it is no possible anymore to directly compare the thresholds found by Smets, De Visser, and Pool [9]. A solution for this would be to ask a handful of the participants to perform the exact same experiment as Smets, De Visser, and Pool [9] to allow for a comparison. This should not be necessary for all participants.

6-4-3 Conditions active experiment

The stimuli levels for the active experiment have been chosen on specific values, as can be seen in Table 6-1 based on the a_1^+ threshold found by Smets, De Visser, and Pool [9]. These are taken as is for now, however, the levels will be tested when the simulator has been configured. For this testing, a test pilot whom has flown many stalls in the Cessna Citation II will test both the settings of the simulator as well as the conditions that have been chosen. Based on this, a final number of conditions and the final a_1 settings will be chosen.

As mentioned in section 3-5, the reason for choosing the method of constant stimuli for this rather than a staircase procedure is because of the uncertainty that the participants themselves might add to the simulated test run which can have a big influence on the staircase outcome. The method of constant stimuli should allow to obtain a psychometric function, which can then be compared to the psychometric function of the passive results. Aiming for either the 50% threshold with the 1D/1U procedure or for the 70.7% point with the staircase

procedure will either way still allow to construct a psychometric function as well, so the method of constant stimuli for the active experiment should give sufficient insight to answer the proposed research question.

6-4-4 Flight Director

If, during testing, it turns out that it is difficult for the test participant to obtain a similar recovery for each test run, it might be an option to add a flight director to the simulation. A flight director is similar to the reference bar shown in Figure 4-1a, but then for the entire recovery procedure. This allows for a more similar recovery procedure for each simulation. As discussed in subsection 5-2-1, this is not expected to be necessary but considered an option when testing shows otherwise.

6-5 Conclusions

This research will investigate to what extent the just noticeable difference thresholds of the Kirchhoff stall model parameters measured during a passive experiment also representative for the thresholds of pilots flying actively in a flight simulation training device. This will be done by setting up a two-fold experiment that will examine the passive and active thresholds of the participants to allow this comparison. The experiment is set up in such a way that the two hypotheses that are formed can be answered. With the results of this experiment, a better understanding of stall modeling for flight simulator training can be achieved, which can ensure that the stall training is done safely and accurately, leading to better pilot responses in case of an actual aircraft stall.

Bibliography

- [1] IATA. *Loss of Control In-Flight Accident Analysis Report*. Tech. rep. International Air Transport Association, 2019.
- [2] ICAO. *Airplane Upset Prevention and Recovery Training Aid, revision 3*. International Civil Aviation Organization, 2017.
- [3] NTSB. *Aircraft Accident Report; Loss of Control on Approach; Colgan Air, Inc.* Washington, DC: National Transportation Safety Board, Feb. 2010.
- [4] BEA. *Final Report on AF447*. 2009. URL: www.bea.aero.
- [5] Dutch Safety Board. *Turkish Airlines Flight TK1951, a Boeing 737-800, Crashed during approach*. 2009. URL: www.safetyboard.nl.
- [6] *Circular Advisory: Stall Prevention and Recovery Training*. Tech. rep. U.S. Department of Transportation, Nov. 2015.
- [7] *Part 60 - FLIGHT SIMULATION TRAINING DEVICE INITIAL AND CONTINUING QUALIFICATION AND USE*. English. Tech. rep. United States Department of Transportation, Oct. 2021. URL: <https://www.ecfr.gov/current/title-14/part-60>.
- [8] EASA. *Certification Specifications for Aeroplane Flight Simulation Training Devices 'CS-FSTD(A)'*. European Aviation Safety Agency, 2018.
- [9] Stephan Smets, Coen C. De Visser, and Daan M. Pool. "Subjective Noticeability of Variations in Quasi-Steady Aerodynamic Stall Dynamics". In: *AIAA Scitech 2019 Forum*. 2019, p. 1485. DOI: [10.2514/6.2019-1485](https://doi.org/10.2514/6.2019-1485).
- [10] Arne Imbrechts, Coen C. De Visser, and Daan M. Pool. "Just Noticeable Differences for Variations in Quasi-Steady Stall Buffet Model Parameters". In: *AIAA SCITECH 2022 Forum*. 2022, p. 0510. DOI: [10.2514/6.2022-0510](https://doi.org/10.2514/6.2022-0510).
- [11] Dirk de Fuijk. "Asymmetric Cessna Citation II Stall Model Identification using a Roll moment-based Kirchhoff method". MSc Thesis. Delft University of Technology, 2023. URL: <http://resolver.tudelft.nl/uuid:6654ffca-9584-42b8-bc0a-ec3138bf789c>.

- [12] Joost Van Ingen. “Dynamic Stall Modeling for the Cessna Citation II”. MSc Thesis. Delft: Delft University of Technology, 2017. URL: <http://resolver.tudelft.nl/uuid:f76439ef-ab2a-4173-8d78-3122a2696e91>.
- [13] A Imbrechts. “Just Noticeable Differences for Variations in Quasi-Steady Stall Buffet Model Parameters”. MSc Thesis. Delft: Delft University of Technology, 2021. URL: <http://resolver.tudelft.nl/uuid:ffb9449a-6477-4216-a71d-c262618e9549>.
- [14] Ed Obert. *Aerodynamic design of transport aircraft*. IOS press, 2009.
- [15] Adson Agrico de Paula, Vitor Gabriel Kleine, and Fabrício De Magalhães Porto. “The thickness effects on symmetrical airfoil flow characteristics at low Reynolds number”. In: *AIAA SciTech Forum - 55th AIAA Aerospace Sciences Meeting*. American Institute of Aeronautics and Astronautics Inc., 2017. ISBN: 9781624104473. DOI: [10.2514/6.2017-1422](https://doi.org/10.2514/6.2017-1422).
- [16] Wolfgang Langewiesche. *Stick and Rudder; An Explanation of the Art of Flying*. McGraw-Hill, Inc, 1944. ISBN: 0-07-036240-8.
- [17] ICAO. *ICAO Doc 10011: MANUAL ON AEROPLANE UPSET PREVENTION AND RECOVERY TRAINING*. International Civil Aviation Organization, 2014.
- [18] John V. Foster, Kevin Cunningham, Charles M. Fremaux, Gautam H. Shah, Eric C. Stewart, Robert A. Rivers, James E. Wilborn, and William Gato. “Dynamics modeling and simulation of large transport airplanes in upset conditions”. In: *Collection of Technical Papers - AIAA Guidance, Navigation, and Control Conference*. Vol. 2. 2005, pp. 826–838. ISBN: 1563477378. DOI: [10.2514/6.2005-5933](https://doi.org/10.2514/6.2005-5933).
- [19] Sven Marschalk, Peter Luteijn, Daan M. Pool, and Coen C. De Visser. “Stall Buffet Modeling using Swept Wing Flight Test Data”. In: *AIAA Scitech 2021 Forum*. 2021, p. 0286. DOI: [10.2514/6.2021-0286](https://doi.org/10.2514/6.2021-0286).
- [20] John N. Ralston, David R. Gingras, Chris Wilkening, and Paul Desrochers. “The application of potential data sources for simulator compliance with ICATEE recommended stall modeling requirements”. In: *AIAA Modeling and Simulation Technologies Conference 2012*. 2012. ISBN: 9781624101830. DOI: [10.2514/6.2012-4568](https://doi.org/10.2514/6.2012-4568).
- [21] D. Fischenberg. “Identification of an Unsteady Aerodynamic Stall Model from Flight Test Data”. In: *20th Atmospheric Flight Mechanics Conference*. American Institute of Aeronautics and Astronautics Inc, 1995, pp. 138–146.
- [22] Eugene A. Morelli, Kevin Cunningham, and Melissa A. Hill. “Global aerodynamic modeling for stall/upset recovery training using efficient piloted flight test techniques”. In: *AIAA Modeling and Simulation Technologies (MST) Conference*. American Institute of Aeronautics and Astronautics Inc., 2013. DOI: [10.2514/6.2013-4976](https://doi.org/10.2514/6.2013-4976).
- [23] Georges Ghazi, Matthieu Bosne, Quentin Sammartano, and Ruxandra Mihaela Botez. “Cessna citation X stall characteristics identification from flight data using neural networks”. In: *AIAA Atmospheric Flight Mechanics Conference, 2017*. American Institute of Aeronautics and Astronautics Inc, AIAA, 2017. ISBN: 9781624104480. DOI: [10.2514/6.2017-0937](https://doi.org/10.2514/6.2017-0937).

- [24] Laurens Van Horssen, Coen C. De Visser, and Daan M. Pool. “Aerodynamic Stall and Buffet Modeling for the Cessna Citation II Based on Flight Test Data”. In: *2018 AIAA Modeling and Simulation Technologies Conference*. 2018, p. 1167. DOI: [10.2514/6.2018-1167](https://doi.org/10.2514/6.2018-1167).
- [25] Joost Van Ingen, Coen C. De Visser, and Daan M. Pool. “Stall Model Identification of a Cessna Citation II from Flight Test Data Using Orthogonal Model Structure Selection”. In: *AIAA Scitech 2021 Forum*. 2021, p. 1725. DOI: [10.2514/6.2021-1725](https://doi.org/10.2514/6.2021-1725).
- [26] L J Van Horssen. “Aerodynamic Stall Modeling for the Cessna Citation II”. MA thesis. Delft: Delft University of Technology, 2016. URL: <https://repository.tudelft.nl>.
- [27] S C E Smets. “Subjective Noticeability of Variations in Quasi-Steady Aerodynamic Stall Dynamics”. MA thesis. 2017. URL: [https://repository.tudelft.nl/..](https://repository.tudelft.nl/)
- [28] P. A. R. Brill. “Optimal Slicing and Partitioning of Flight Data for Highest Reliability of Stall Model Parameter Estimates”. Preliminary Thesis Report. 2022.
- [29] *Merriam-Webster Dictionary: Hysteresis*. accessed [01/03/2023]. URL: <https://www.merriam-webster.com/dictionary/hysteresis>.
- [30] J Dennis White and Judy A Vilmain. “The psychophysics of price: A critique of the Weber-Fechner approach in consumer behavior”. In: *Proceedings of the 1986 Academy of Marketing Science (AMS) Annual Conference*. Springer. 1986, pp. 30–34.
- [31] Anthony Adams, C Anthony Di Benedetto, and Rajan Chandran. “Can you reduce your package size without damaging sales?” In: *Long Range Planning* 24.4 (1991), pp. 86–96.
- [32] Gustav Theodor Fechner. *Elemente der Psychophysik*. Brietkopf und Härtel, 1860.
- [33] S.S. Stevens. *Psychophysics: Introduction to its perceptual, neural, and social prospects*. John Wiley & Sons, Inc., 1975.
- [34] William A Simpson. “The method of constant stimuli is efficient”. In: *Perception & psychophysics* 44 (1988), pp. 433–436.
- [35] Simon Grondin. *Psychology of perception*. Springer, 2016.
- [36] HCCH Levitt. “Transformed up-down methods in psychoacoustics”. In: *The Journal of the Acoustical society of America* 49.2B (1971), pp. 467–477.
- [37] Niel A. Macmillan and C. Douglas Creelman. *Detection Theory: A User’s Guide*. Cambridge University Press, 1991.
- [38] Tom N. Cornsweet. “The Staircase-Method in Psychophysics”. In: *The American Journal of Psychology* 75.3 (1962), pp. 485–491. ISSN: 00029556. URL: <http://www.jstor.org/stable/1419876> (visited on 04/19/2023).
- [39] Frederick A.A. Kingdom and Nicolaas Prins. *Psychophysics: A practical introduction*. second. Mica Haley, 2016.
- [40] Marjorie R Leek. “Adaptive procedures in psychophysical research”. In: *Perception & Psychophysics* 63.8 (2001), pp. 1279–1292.
- [41] M M Taylor and C Douglas Creelman. “PEST: Efficient estimates on probability functions”. In: *The Journal of the Acoustical Society of America* 41.4A (1967), pp. 782–787.

- [42] Alex Pentland. “Maximum likelihood estimation: The best PEST”. In: 28.4 (1980), pp. 377–379.
- [43] R. J. A. W. Hosman and J. C. van der Vaart. *Vestibular Models and Thresholds of Motion Perception. Results of tests in a flight simulator*. Tech. rep. LR-265. Delft: Delft University of Technology, 1978. URL: <http://resolver.tudelft.nl/uuid:72bb1e55-7304-459f-a47c-dae7984418e3>.
- [44] A. R. Valente Pais, D. M. Pool, A. M. De Vroome, M. M. Van Paassen, and M. Mulder. “Pitch motion perception thresholds during passive and active tasks”. In: *Journal of Guidance, Control, and Dynamics* 35.3 (2012), pp. 904–918. ISSN: 15333884. DOI: [10.2514/1.54987](https://doi.org/10.2514/1.54987).
- [45] Jeffery A Schroeder, Judith S Burki-Cohen, David Shikany, David R Gingras, and Paul P Desrochers. “An evaluation of several stall models for commercial transport training”. In: *AIAA Modeling and Simulation Technologies Conference*. 2014, p. 1002.
- [46] Kevin Cunningham, Gautam H. Shah, Patrick C. Murphy, Melissa A. Hill, and Brent P. Pickering. “Pilot sensitivity to simulator flight dynamics model formulation for stall training”. In: *AIAA Scitech 2019 Forum*. American Institute of Aeronautics and Astronautics Inc, AIAA, 2019. ISBN: 9781624105784. DOI: [10.2514/6.2019-0717](https://doi.org/10.2514/6.2019-0717).
- [47] Peter R. Grant, Gregory J. Moszczynski, and Jeffery A. Schroeder. “Post-stall flight model fidelity effects on full stall recovery training”. In: *2018 Modeling and Simulation Technologies Conference*. American Institute of Aeronautics and Astronautics Inc, AIAA, 2018. ISBN: 9781624105517. DOI: [10.2514/6.2018-2937](https://doi.org/10.2514/6.2018-2937).
- [48] Euclid C Holleman and David L Boslaugh. *A simulator investigation of factors affecting the design and utilization of a stick pusher for the prevention of airplane pitch-up*. Tech. rep. 1958.
- [49] Skybrary. *Deep Stall*. <https://www.skybrary.aero/articles/deep-stall#:~:text=Definition,normal%20stall%20recovery%20actions%20ineffective>. [accessed: 16-05-2023].
- [50] Thomas Lombaerts, Stefan Schuet, Vahram Stepanyan, John Kaneshige, Gordon Hardy, Kimberlee Shish, and Peter Robinson. “Piloted simulator evaluation results of flight physics based stall recovery guidance”. In: *AIAA Guidance, Navigation, and Control Conference, 2018*. American Institute of Aeronautics and Astronautics Inc, AIAA, Jan. 2018. ISBN: 9781624105265. DOI: [10.2514/6.2018-0383](https://doi.org/10.2514/6.2018-0383).
- [51] Vahram Stepanyan, Kalmanje S Krishnakumar, John Kaneshige, and Diana M Acosta. “Stall recovery guidance algorithms based on constrained control approaches”. In: *AIAA Guidance, Navigation, and Control Conference*. 2016, p. 0878.
- [52] Commission D’Enquete sur les Accidents et Incidents d’Aviation Civile. *Final Report; Accident on 24 July 2014 near Gossi (Mali)*. Republique du Mali: Ministere de L’Equipement, des Transports et du Desenclavement, Apr. 2016.
- [53] Stefan Schuet, Thomas Lombaerts, Vahram Stepanyan, John Kaneshige, Kimberlee Shish, Peter Robinson, and Gordon H Hardy. *Vertical Motion Simulator Experiment on Stall Recovery Guidance*. Tech. rep. 2017.

- [54] European Union Aviation Safety Agency. *Translation of the French DGAC leaflet on stalls*. accessed [12/04/2023]. Nov. 2012. URL: <https://www.easa.europa.eu/en/downloads/43430/en>.
- [55] *Aircraft Stall Model Identification*. [accessed: 31/07/2023]. Aug. 2015. URL: <https://cs.lr.tudelft.nl/citation/projects/stallmodels/>.
- [56] Wietse D. Ledegang and Eric L. Groen. “Stall Recovery in a Centrifuge-Based Flight Simulator With an Extended Aerodynamic Model”. In: *International Journal of Aviation Psychology* 25.2 (Apr. 2015), pp. 122–140. ISSN: 10508414. DOI: [10.1080/10508414.2015.1131085](https://doi.org/10.1080/10508414.2015.1131085).
- [57] M.A. Van den Hoek, C.C. De Visser, and D.M. Pool. “Identification of a Cessna Citation II model based on flight test data”. In: *Advances in Aerospace Guidance, Navigation and Control: Selected Papers of the Fourth CEAS Specialist Conference on Guidance, Navigation and Control Held in Warsaw, Poland, April 2017*. Springer. 2018, pp. 259–277.
- [58] J.A. Mulder, W.H.J.J. van Staveren, J.C. van der Vaart, E. de Weerdt, C.C. de Visser, A.C. in ’t Veld, and E. Mooij. *Flight Dynamics, Lecture Notes AE3202*. Delft University of Technology, 2013.
- [59] Dies Vuik. “Effect of external vehicle perturbations on tactile perception by electrovibration”. MSc Thesis. Delft: Delft University of Technology, 2023. URL: <http://resolver.tudelft.nl/uuid:bcaac32b-f789-4830-bdcc-943290ea515d>.

Part III

Appendices to Preliminary Thesis Report

Appendix A

Additional Figures

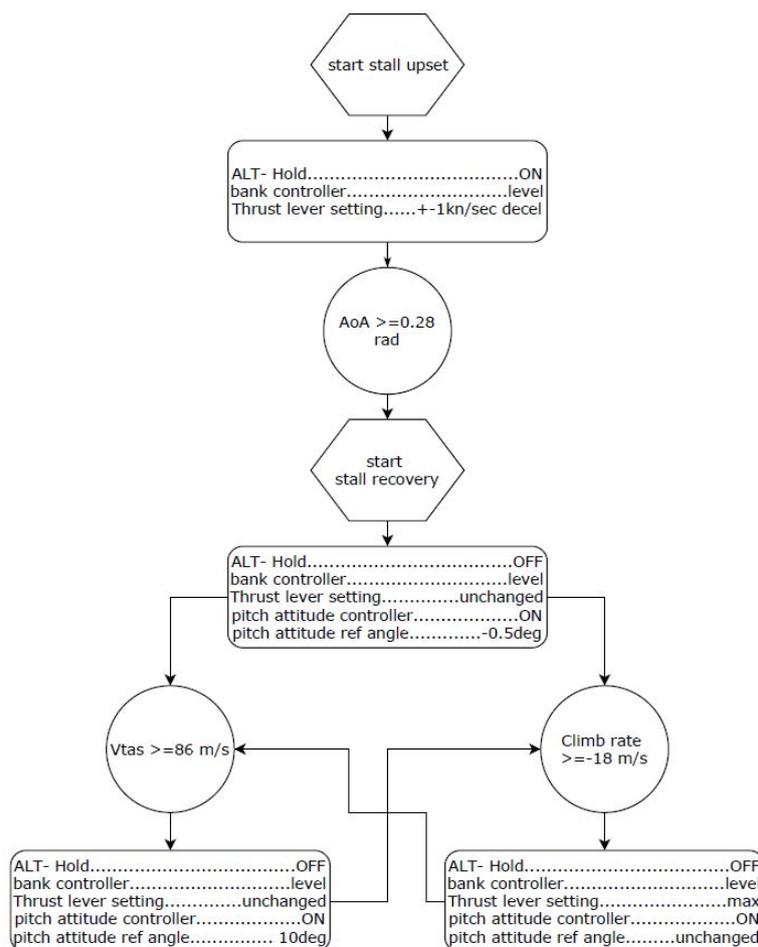


Figure A-1: Stall autopilot flow chart with steps during recovery phase [9]

A-1 Sensitivity analysis part 1

A-1-1 States for different simulations centered around the stall

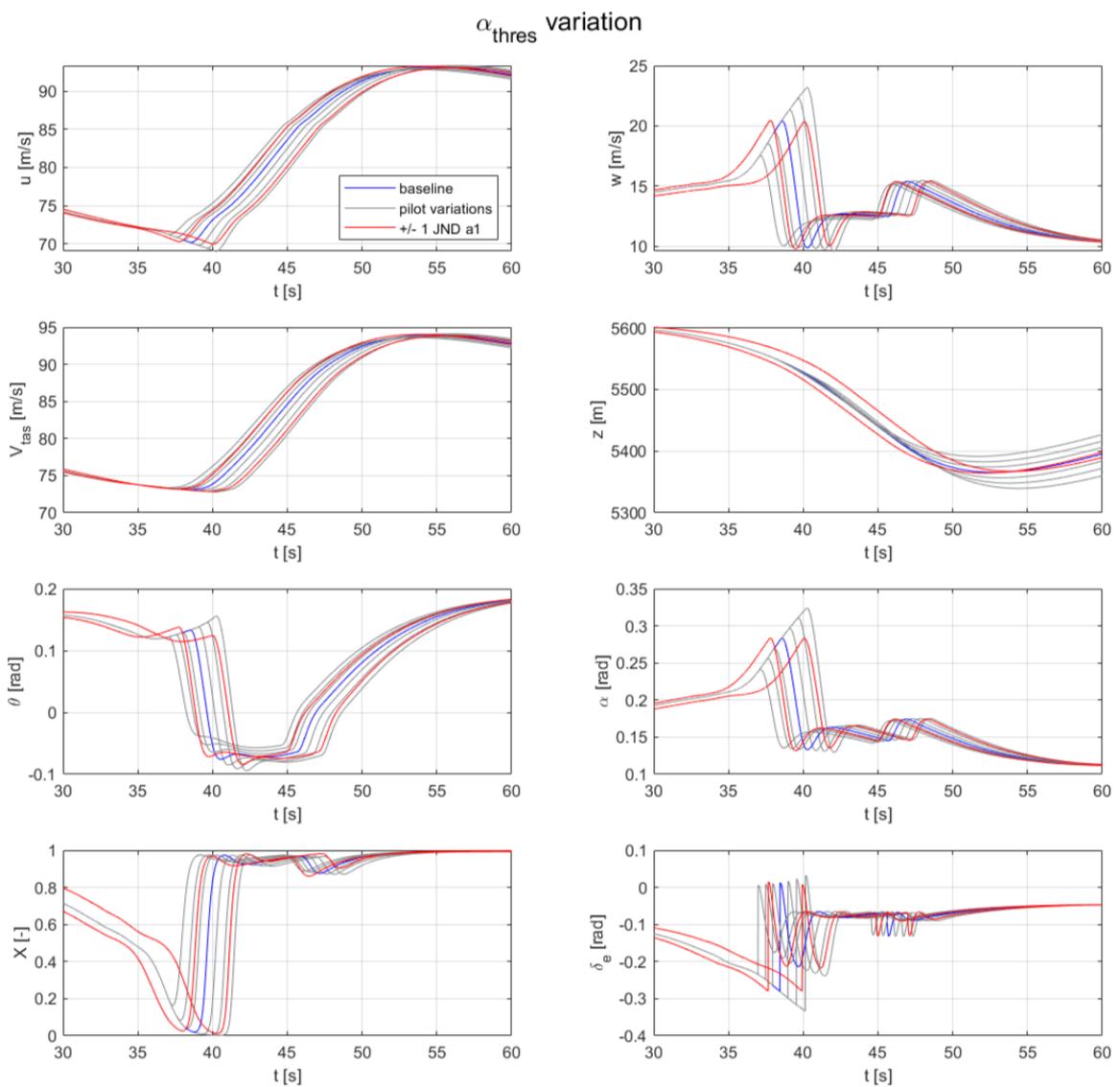


Figure A-2: Influence of different settings of $\alpha_{\text{threshold}}$

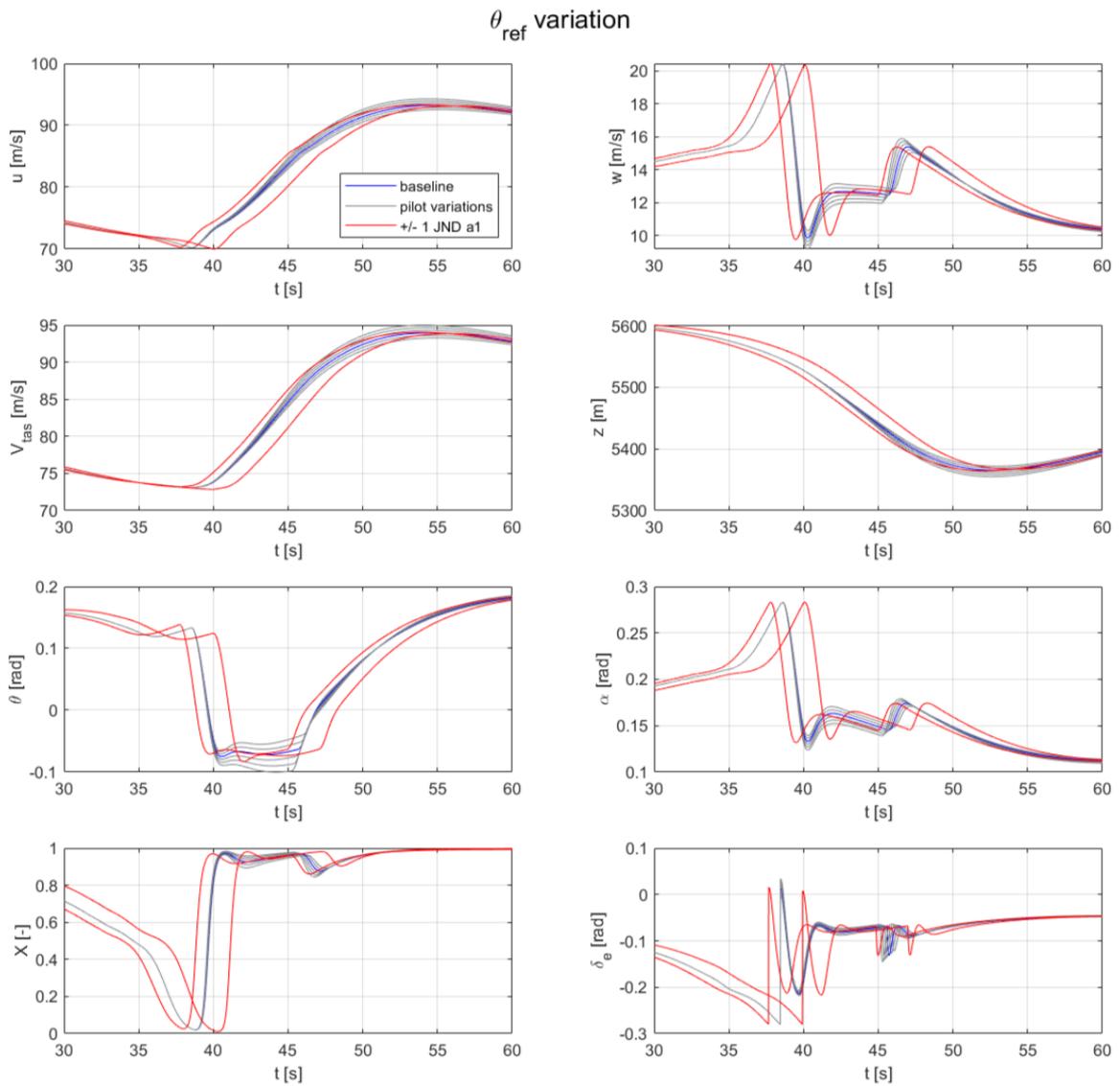


Figure A-3: Influence of different settings of θ_{ref}

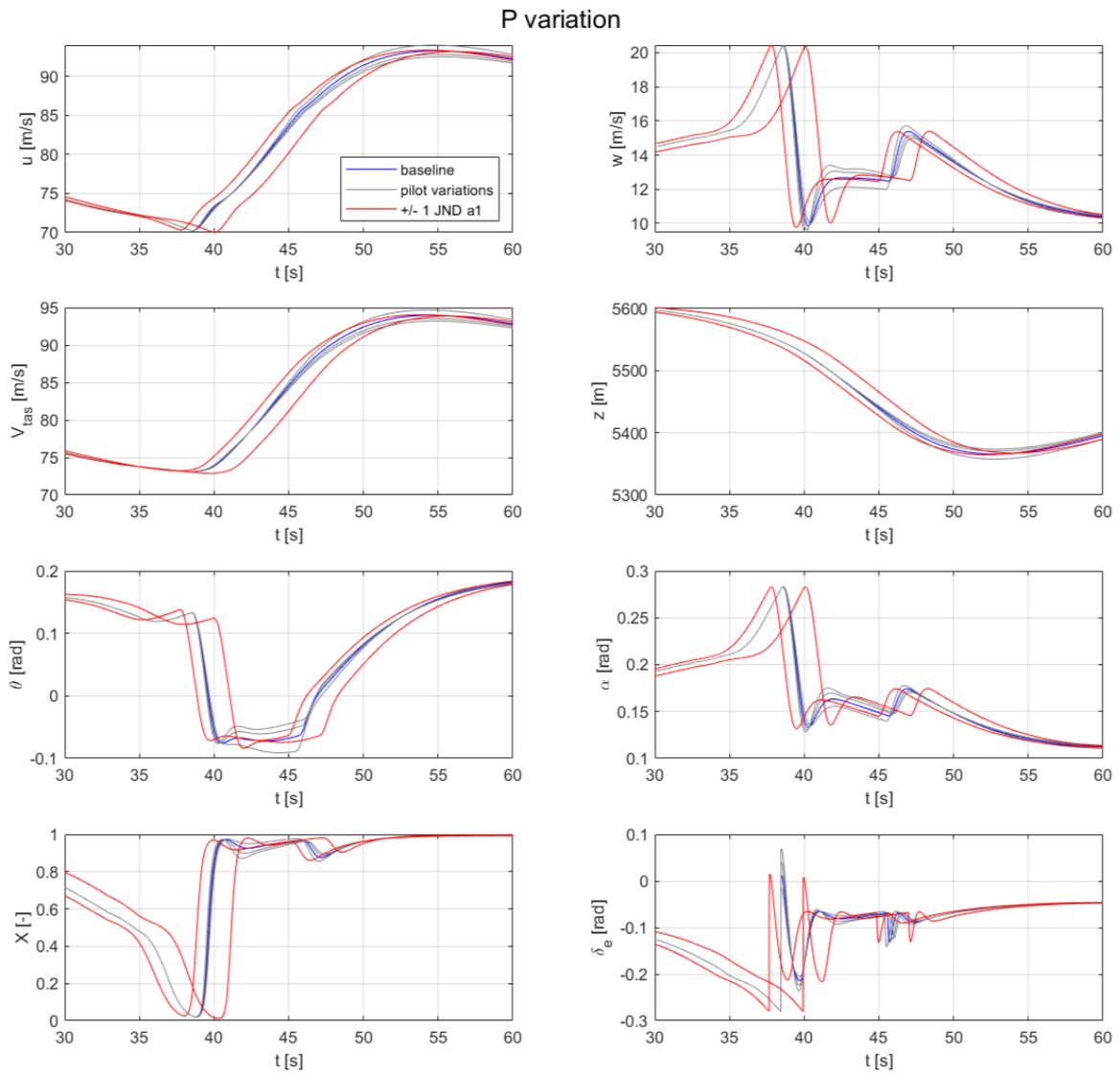


Figure A-4: Influence of different settings of gain P_θ

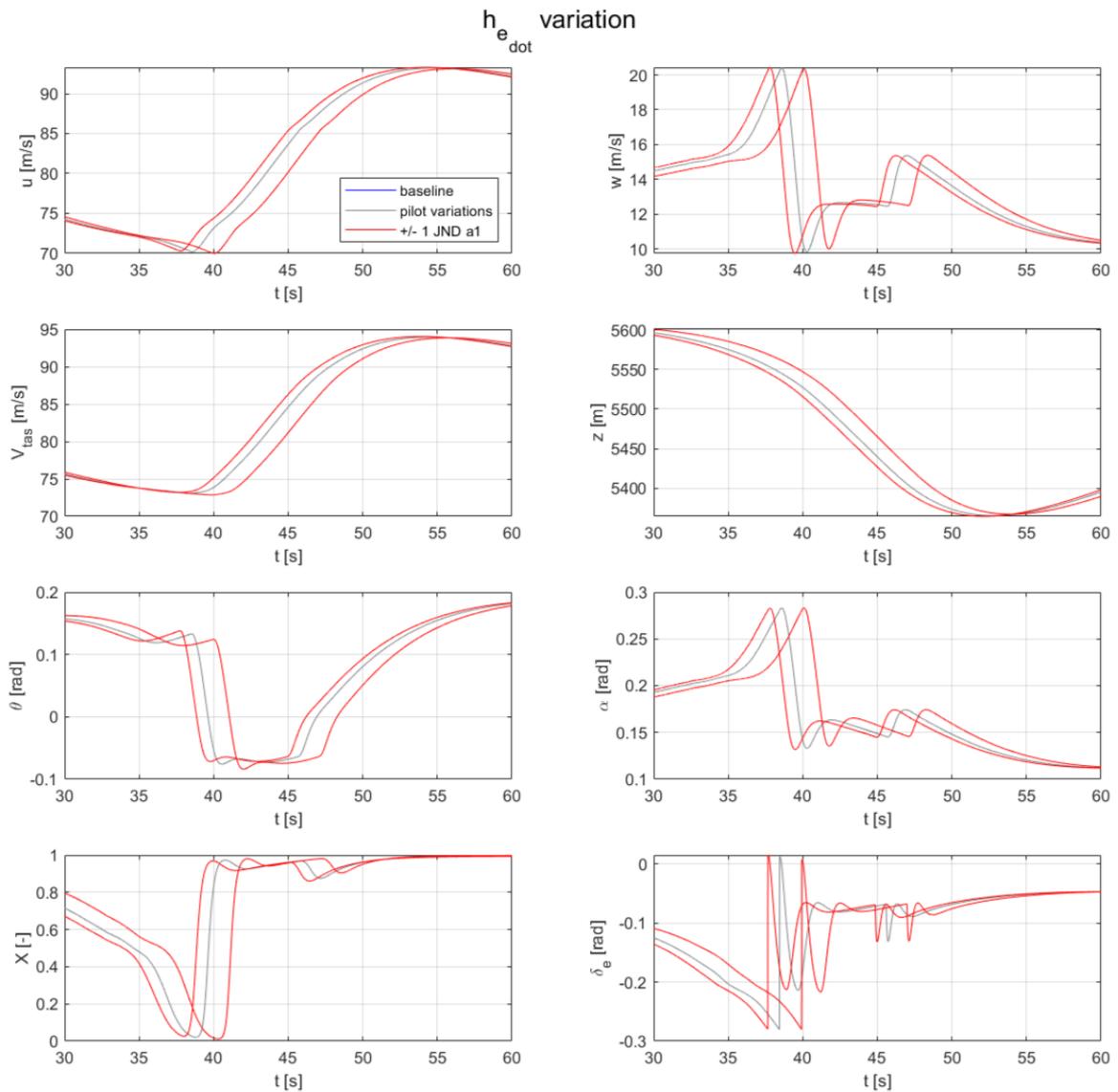


Figure A-5: Influence of different settings of \dot{h}_e

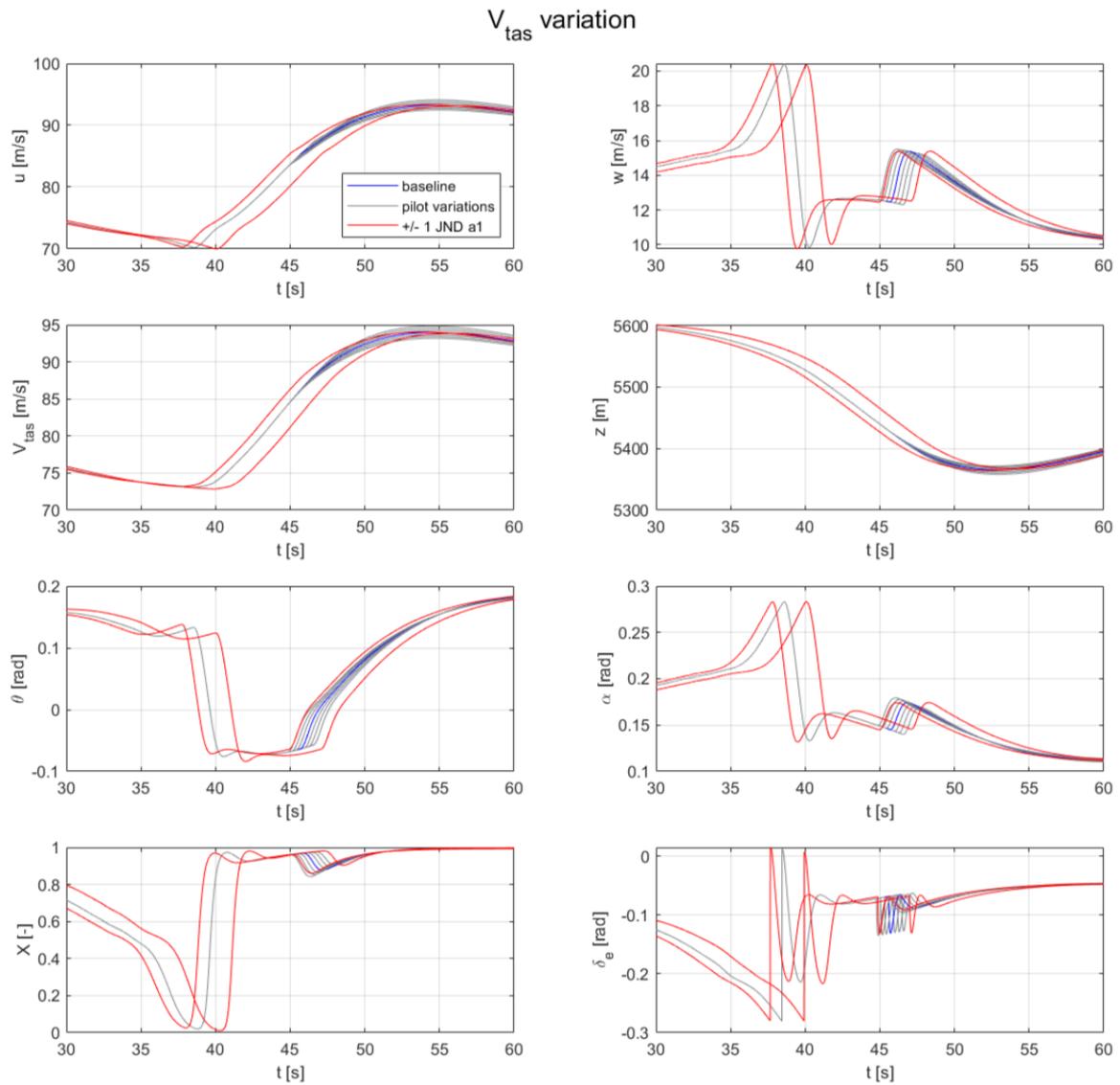


Figure A-6: Influence of different settings of V_{TAS}

A-1-2 Delta of the states for different simulations centered around the stall with respect to the baseline

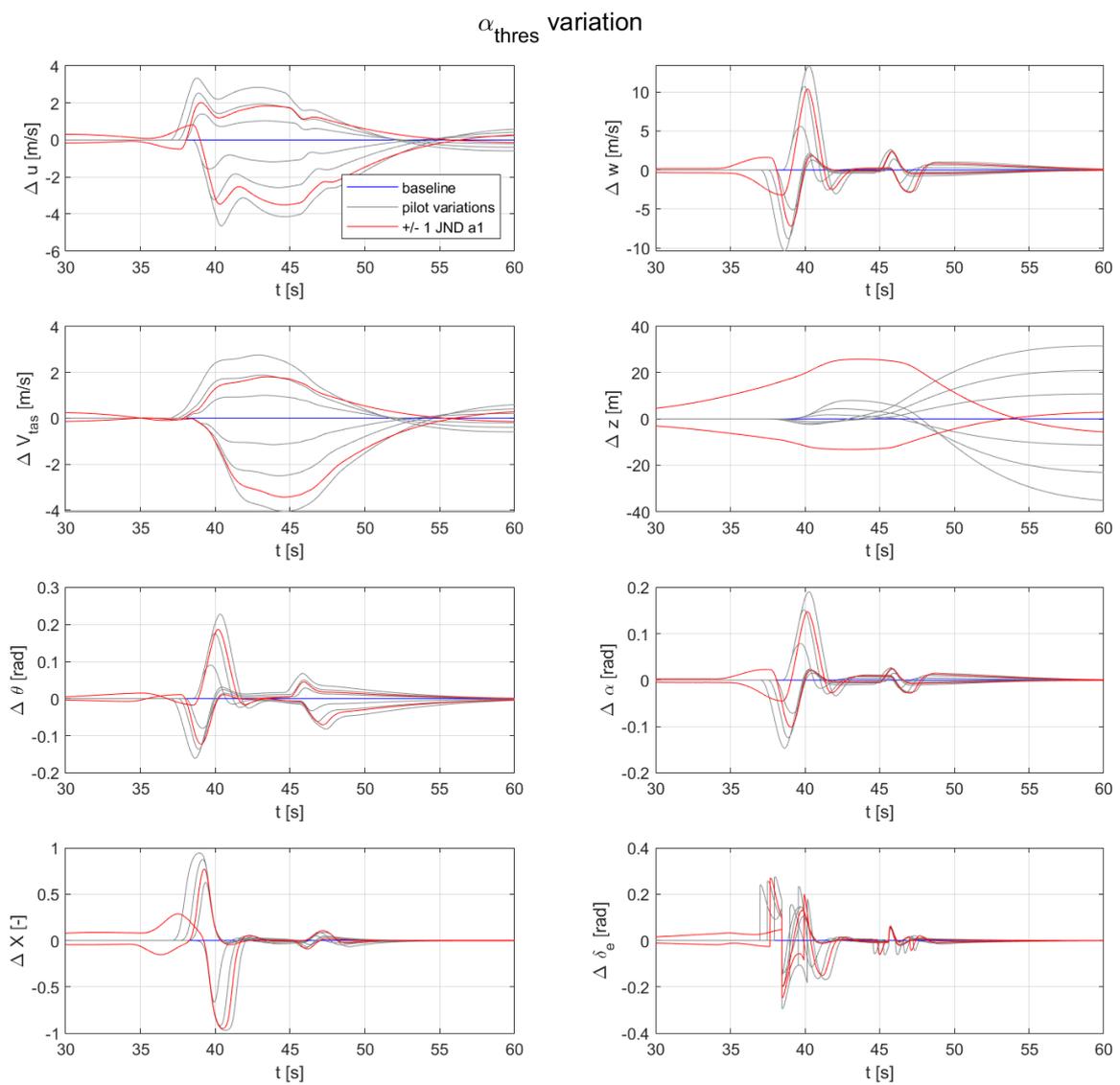


Figure A-7: Delta states of different settings of $\alpha_{threshold}$ with respect to the baseline

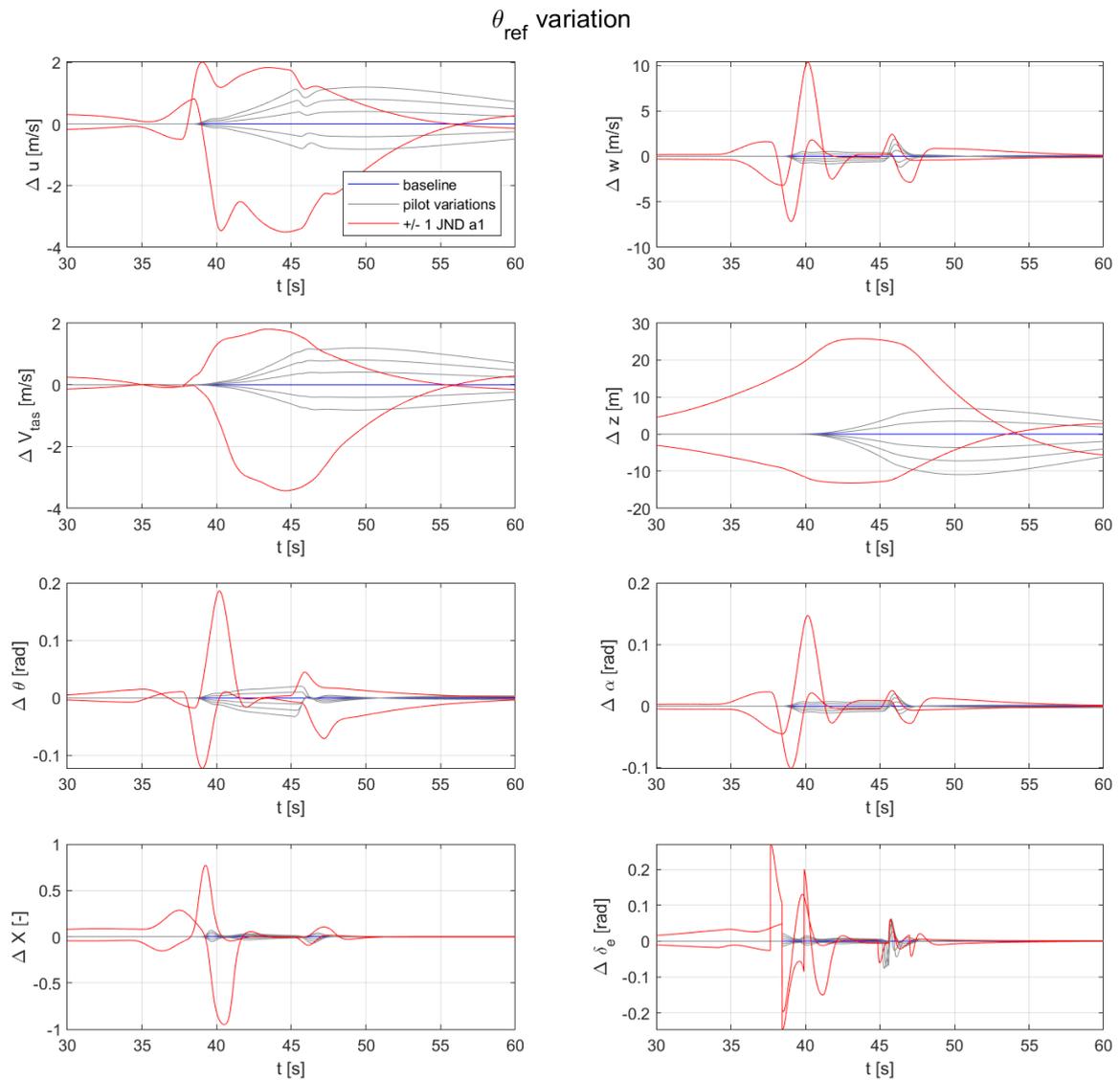


Figure A-8: Delta states of different settings of θ_{ref} with respect to the baseline

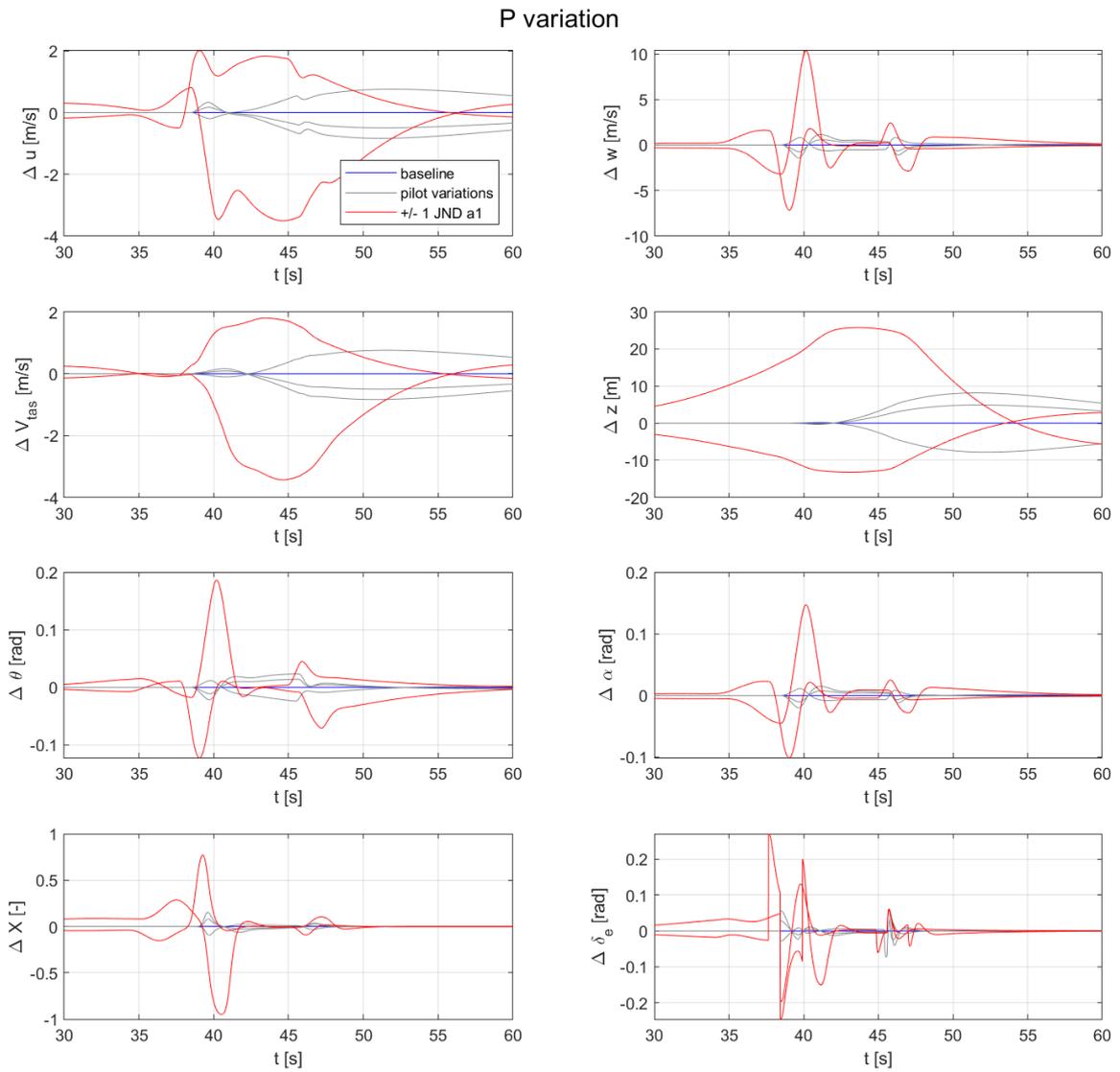


Figure A-9: Delta states of different settings of gain P_θ with respect to the baseline

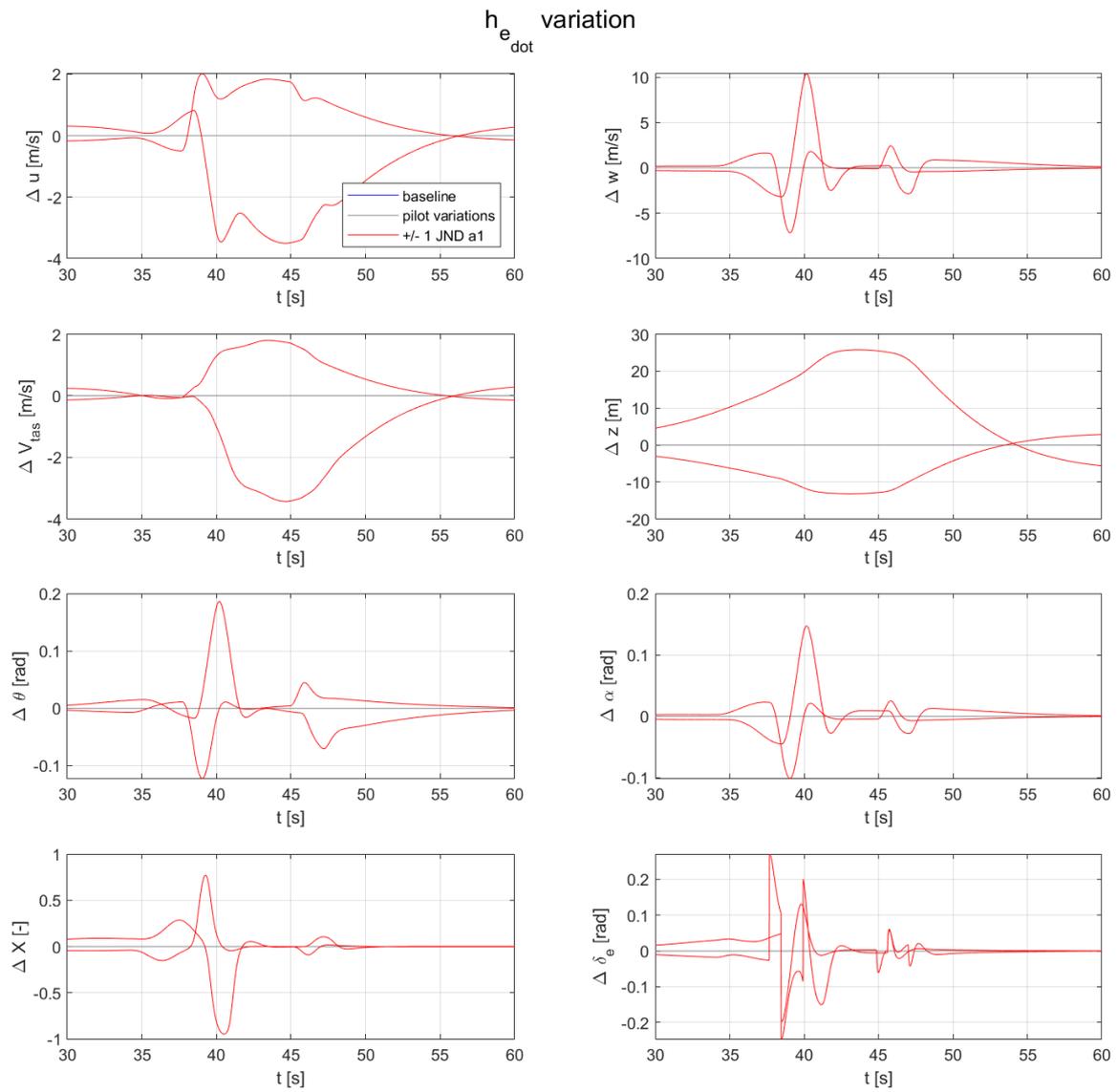


Figure A-10: Delta states of different settings of \dot{h}_e with respect to the baseline

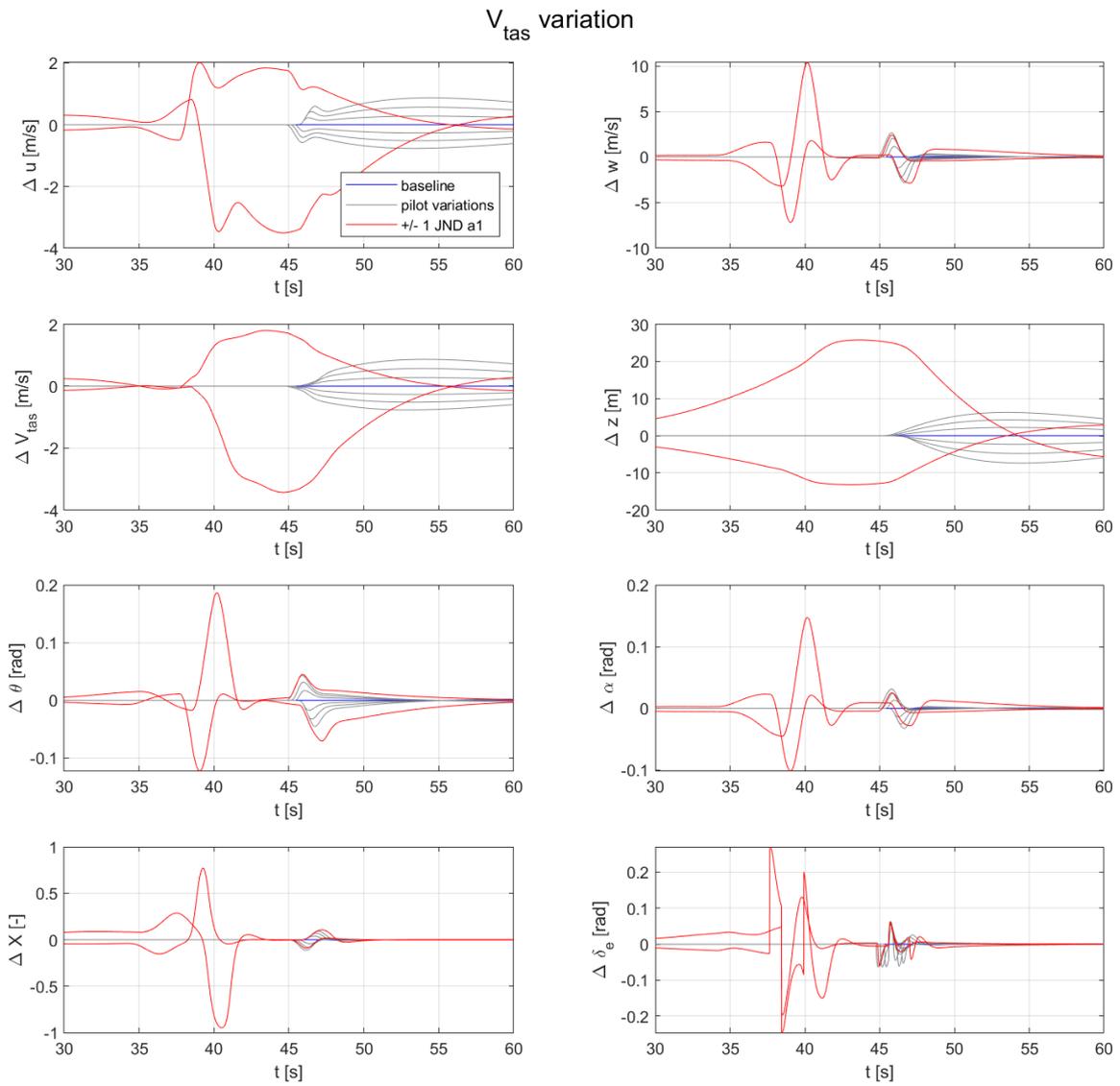


Figure A-11: Delta states of different settings of V_{TAS} with respect to the baseline

A-2 Sensitivity analysis part 2

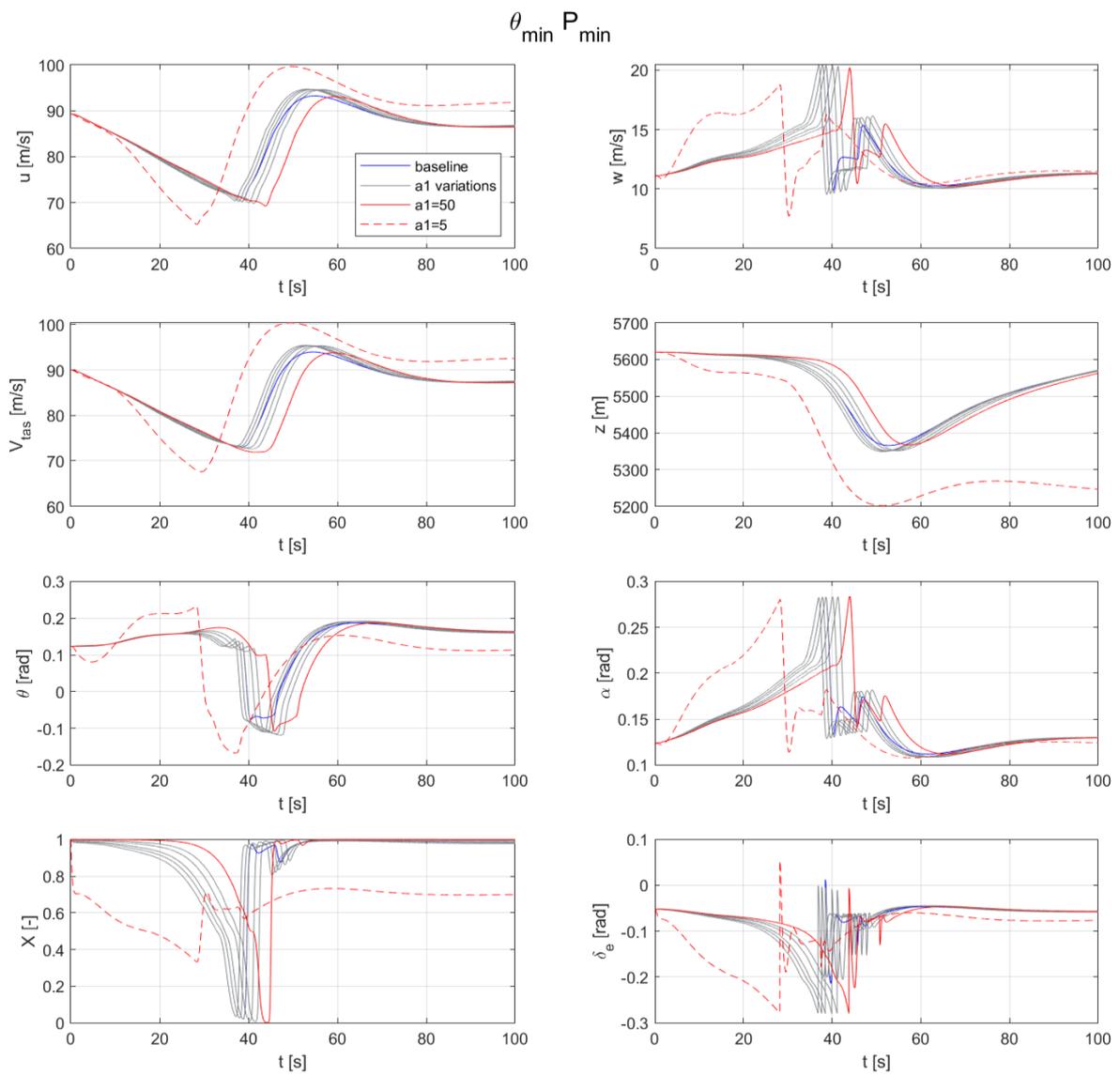


Figure A-12: Different a_1 settings with $\theta_{ref} = -3.5$ and gain $P_\theta = 0.2$

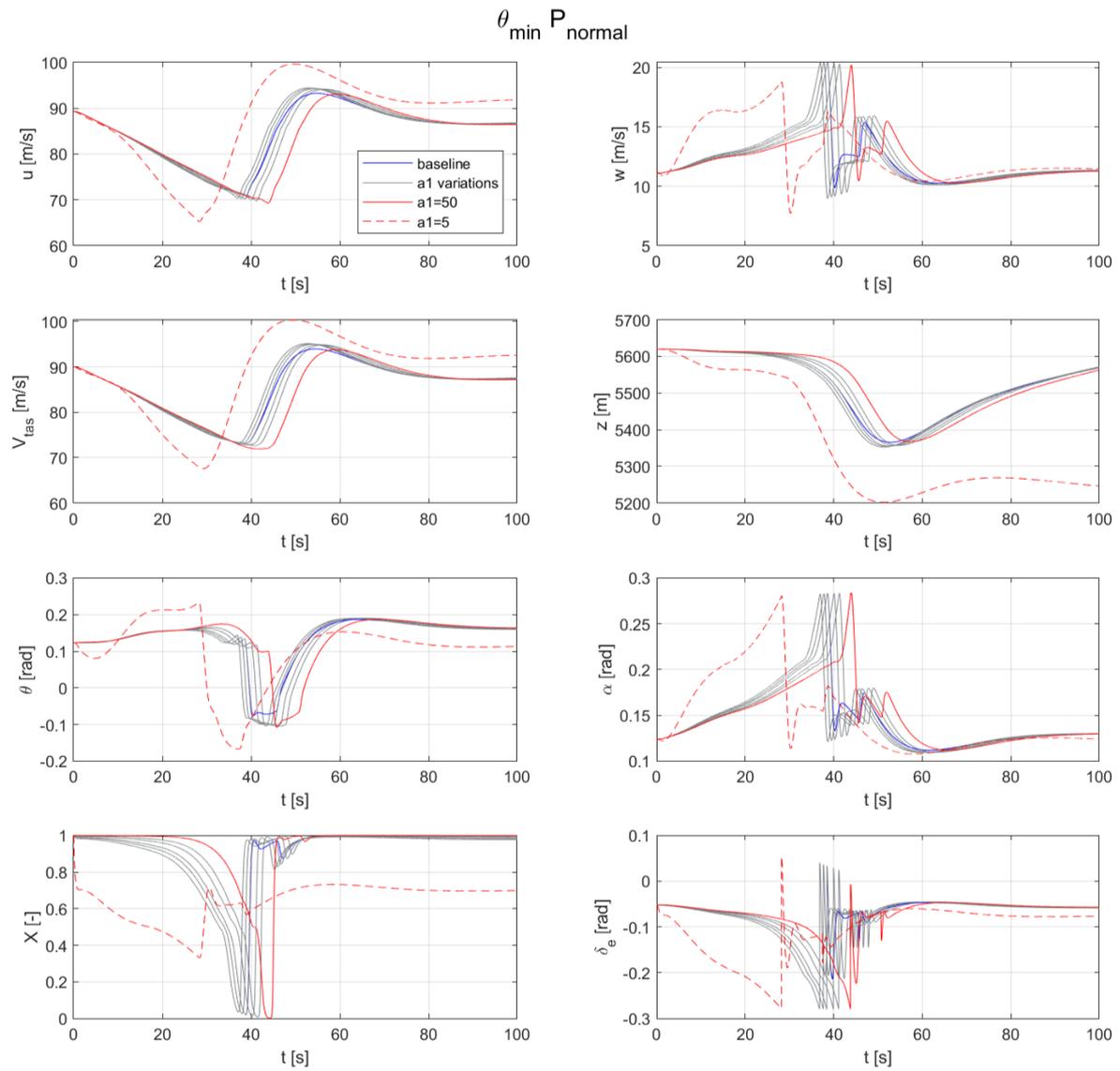


Figure A-13: Different a_1 settings with $\theta_{ref} = -3.5$ and gain $P_\theta = 0.2$

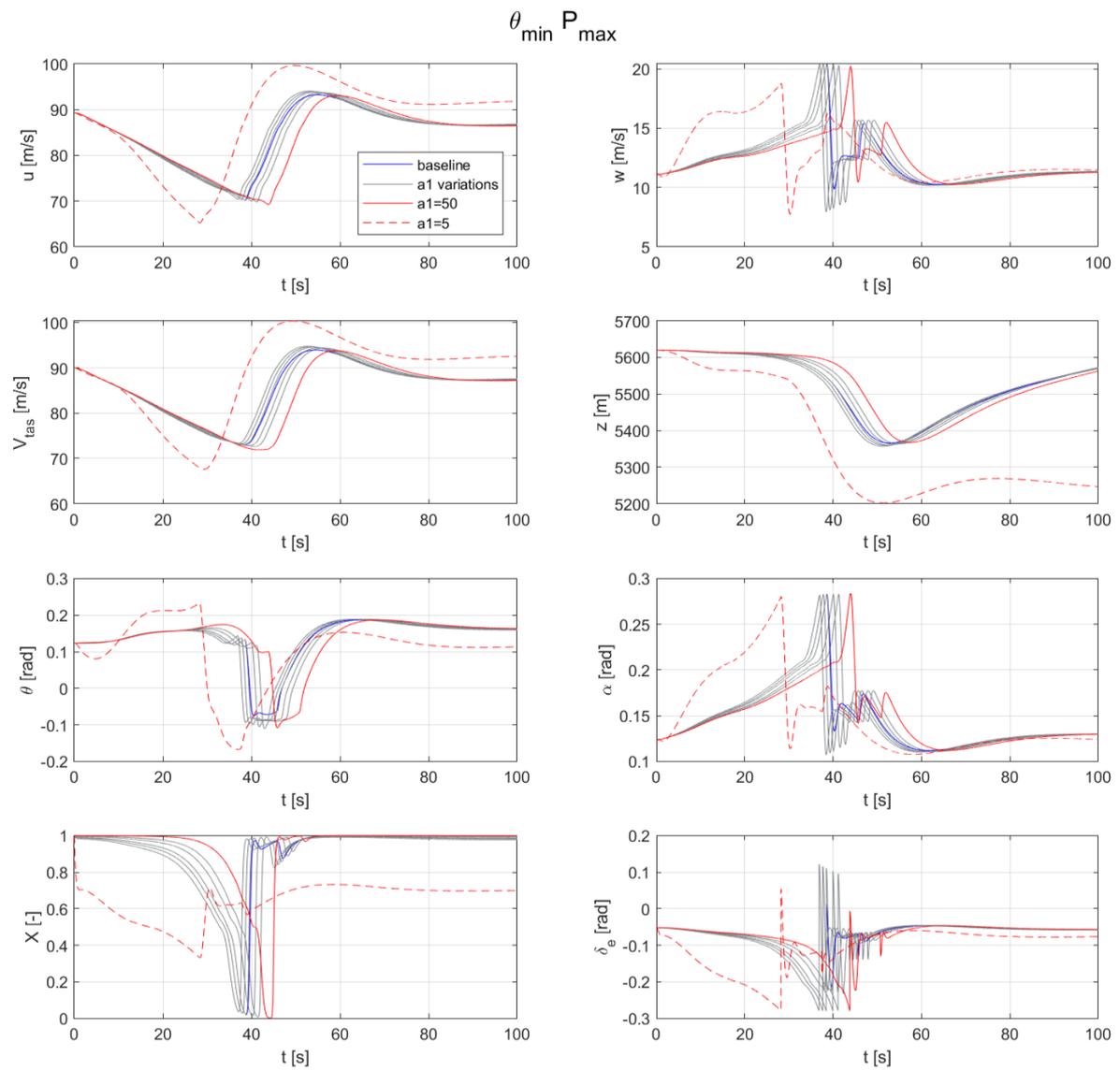


Figure A-14: Different a_1 settings with $\theta_{ref} = -3.5$ and gain $P_\theta = 0.2$

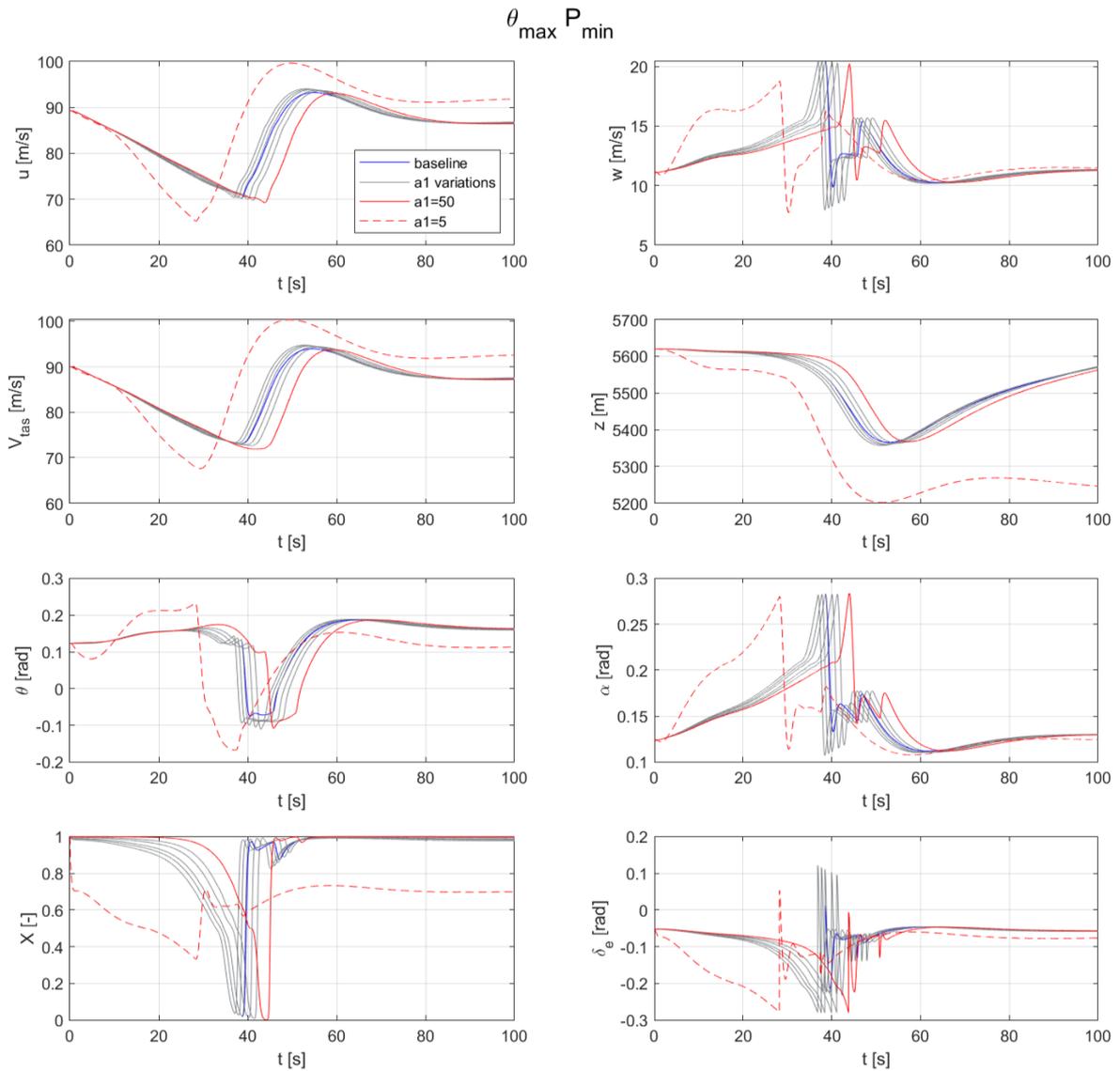


Figure A-15: Different a_1 settings with $\theta_{ref} = 1.5$ and gain $P_\theta = 0.2$

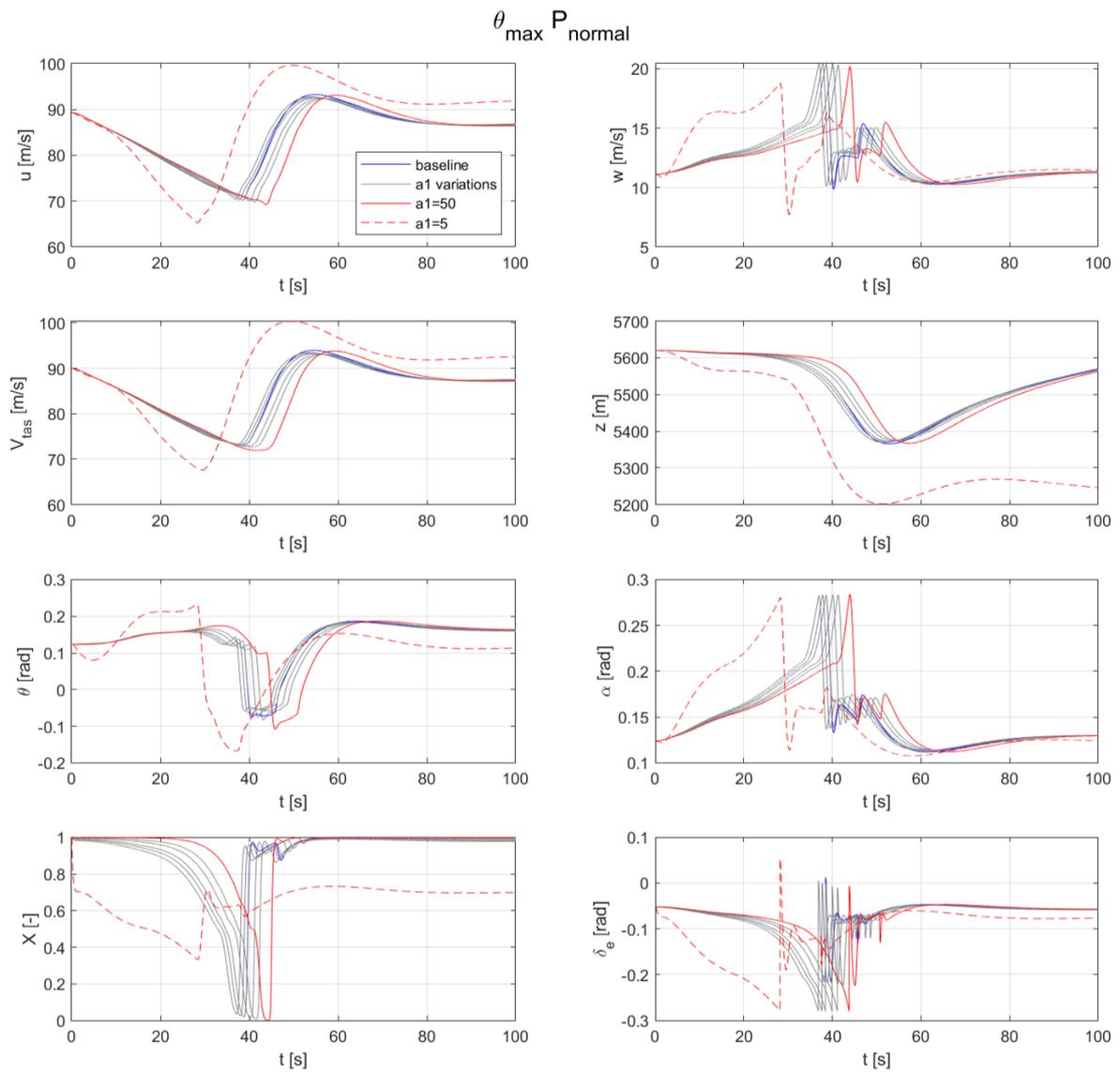


Figure A-16: Different a_1 settings with $\theta_{ref} = 1.5$ and gain $P_\theta = 0.4$

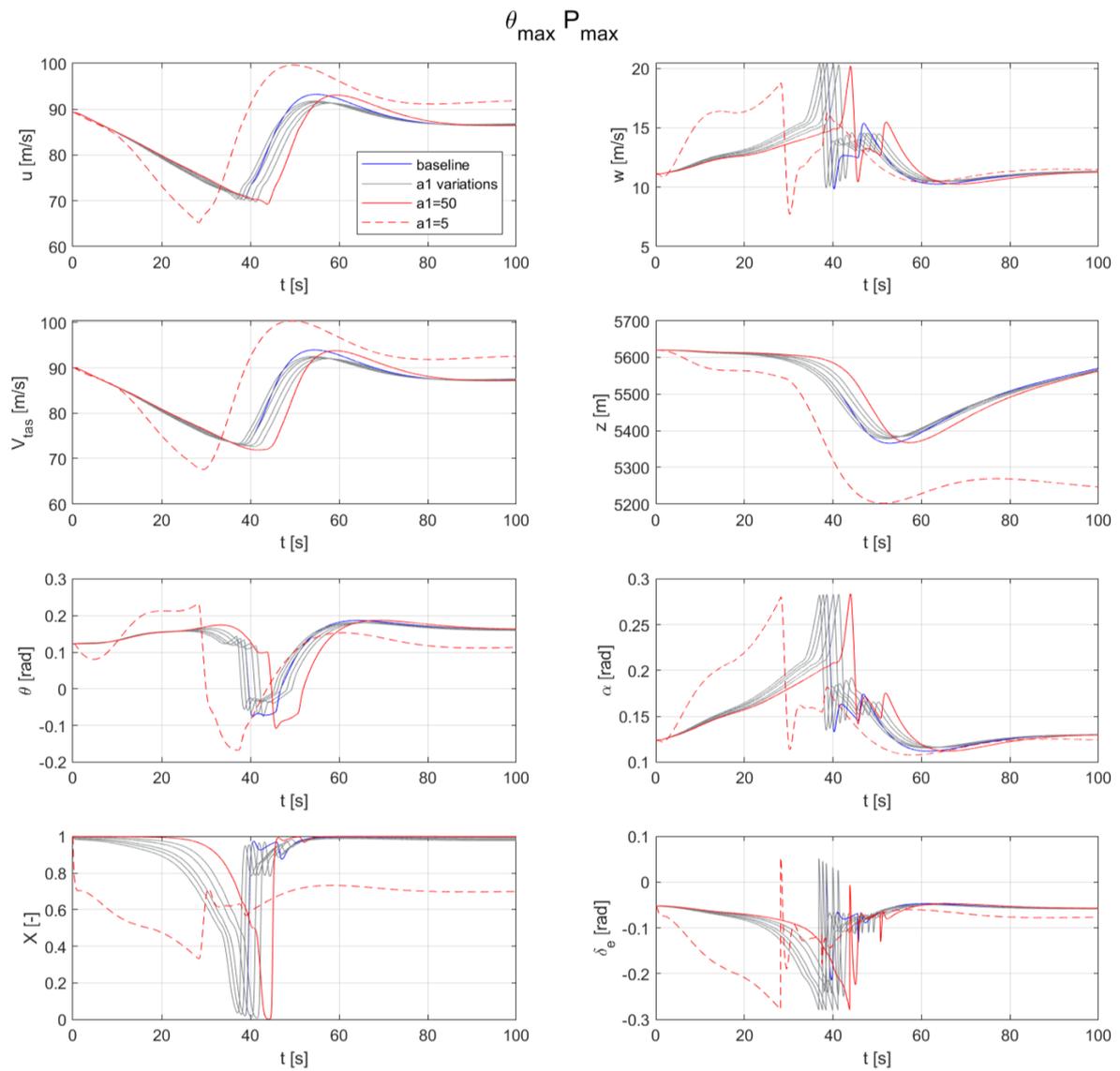


Figure A-17: Different a_1 settings with $\theta_{ref} = 1.5$ and gain $P_\theta = 0.8$

A-3 Experiment Design

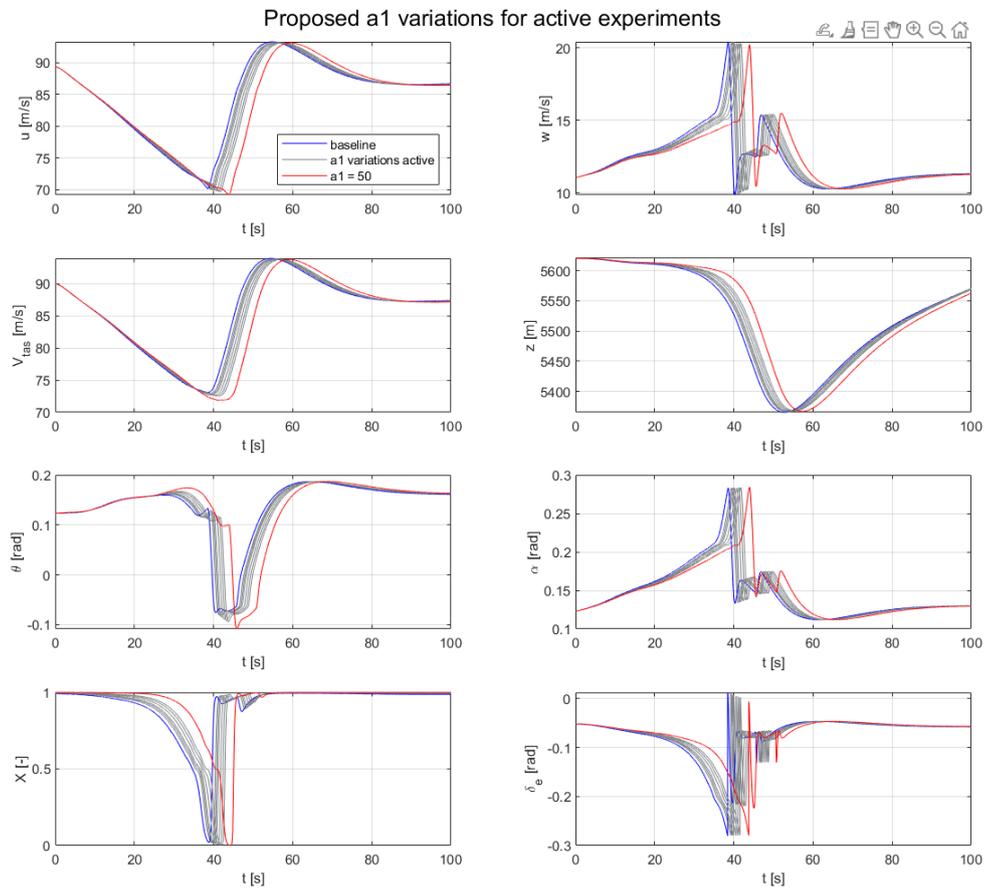


Figure A-18: Different a_1 settings of the active experiment, and the maximum a_1 setting from the passive experiment

Appendix B

Additional Tables

B-1 Sensitivity Analysis

Table B-1: VAF with respect to $\dot{h}_e = -18$ [m/s]

| | \dot{h}_e [m/s] | | | |
|----------------|-------------------|------------|------------|------------|
| | -22 | -20 | -16 | -14 |
| q | 100 | 100 | 100 | 100 |
| u | 100 | 100 | 100 | 100 |
| v | 100 | 100 | 100 | 100 |
| w | 100 | 100 | 100 | 100 |
| θ | 100 | 100 | 100 | 100 |
| x | 100 | 100 | 100 | 100 |
| y | 100 | 100 | 100 | 100 |
| z | 100 | 100 | 100 | 100 |
| X | 100 | 100 | 100 | 100 |
| δ_e | 100 | 100 | 100 | 100 |
| α | 100 | 100 | 100 | 100 |
| Average | 100 | 100 | 100 | 100 |

B-2 Sensitivity Analysis part 2

B-2-1 Sensitivity with minimal gain $P_{\theta} = 0.2$

Scenario 1

Table B-2: Sensitivity of τ_1 with minimum θ_{ref} and minimum gain P_{θ}

| | $\theta_{\text{ref}} = -3.5[\text{deg}]$ | | | |
|----------------|--|-----------------------|-----------------------|-----------------------|
| | -2JND τ_1 | -1JND τ_1 | +1JND τ_1 | +2JND τ_1 |
| q | 99.9975 | 99.9967 | 99.9992 | 99.9946 |
| u | 99.3736 | 99.3004 | 99.3422 | 99.1962 |
| v | 99.3251 | 99.2608 | 99.2600 | 99.0629 |
| w | 97.1061 | 95.2392 | 96.2107 | 92.7171 |
| θ | 98.6572 | 98.4274 | 97.4493 | 95.7236 |
| x | 99.9975 | 99.9972 | 99.9980 | 99.9982 |
| y | 99.9967 | 99.9975 | 99.9992 | 99.9946 |
| z | 99.7252 | 99.7163 | 99.5552 | 99.3547 |
| X | 98.6988 | 95.2005 | 97.7415 | 93.4867 |
| δ_e | 90.0319 | 82.7626 | 89.3133 | 80.9111 |
| α | 98.5151 | 97.3379 | 98.0283 | 95.9505 |
| Average | 97.0216 | 98.3113 | 97.8996 | 96.0358 |

Scenario 2

99.9599570473076 99.9939316717543 99.9967341231723 99.9974496960982 99.9871663587847

Table B-3: Sensitivity of a_1 with minimum θ_{ref} and minimum gain P_{θ}

| | $\theta_{\text{ref}} = -3.5[\text{deg}]$ | | | |
|----------------|--|--------------------|--------------------|--------------------|
| | -2JND a_1 | -1JND a_1 | +1JND a_1 | +2JND a_1 |
| q | 99.9871 | 99.9974 | 99.9600 | 99.9939 |
| u | 95.9441 | 98.3824 | 97.7091 | 93.0707 |
| v | 95.8070 | 98.3414 | 97.2518 | 92.1094 |
| w | 50.0606 | 81.3322 | 55.1267 | 32.6204 |
| θ | 85.2893 | 94.9593 | 82.9799 | 64.0301 |
| x | 99.9910 | 99.9952 | 99.9991 | 99.9966 |
| y | 99.9872 | 99.9974 | 99.9600 | 99.9939 |
| z | 98.5701 | 99.4163 | 98.6084 | 96.0055 |
| X | 59.3049 | 86.0669 | 60.3193 | 20.9601 |
| δ_e | 65.1083 | 68.2787 | 66.6370 | 49.9489 |
| α | 69.1772 | 88.7801 | 73.1083 | 58.9857 |
| Average | 83.5664 | 92.3223 | 84.6999 | 73.4289 |

Scenario 3**Table B-4:** Sensitivity of τ_1 with maximum θ_{ref} and minimum gain P_θ

| | $\theta_{\text{ref}} = 1.5[\text{deg}]$ | | | |
|----------------|---|-----------------------|-----------------------|-----------------------|
| | -2JND τ_1 | -1JND τ_1 | +1JND τ_1 | +2JND τ_1 |
| q | 99.9999 | 99.9999 | 99.9975 | 99.9853 |
| u | 99.9156 | 99.9720 | 99.8931 | 99.7161 |
| v | 99.9146 | 99.9707 | 99.8765 | 99.6534 |
| w | 98.1099 | 99.5586 | 97.6560 | 93.8087 |
| θ | 99.4618 | 99.8876 | 99.1296 | 97.6609 |
| x | 99.9998 | 99.9999 | 99.9999 | 99.9999 |
| y | 99.9999 | 99.9999 | 99.9975 | 99.9853 |
| z | 99.9395 | 99.9860 | 99.9063 | 99.7501 |
| X | 96.5738 | 99.4453 | 97.1521 | 92.5047 |
| δ_e | 85.1050 | 92.1397 | 90.1275 | 80.5520 |
| α | 98.8454 | 99.7393 | 98.5987 | 96.2794 |
| Average | 97.9877 | 99.1544 | 98.3943 | 96.3555 |

Scenario 4**Table B-5:** Sensitivity of a_1 with maximum θ_{ref} and minimum gain P_θ

| | $\theta_{\text{ref}} = 1.5[\text{deg}]$ | | | |
|----------------|---|--------------------|--------------------|--------------------|
| | -2JND a_1 | -1JND a_1 | +1JND a_1 | +2JND a_1 |
| q | 99.9897 | 99.9997 | 99.9593 | 99.9991 |
| u | 97.5636 | 99.4100 | 97.6944 | 92.8610 |
| v | 97.4824 | 99.4266 | 97.2989 | 92.0676 |
| w | 57.5819 | 86.3568 | 56.6889 | 36.5780 |
| θ | 87.6760 | 96.6194 | 85.9001 | 69.4101 |
| x | 99.9967 | 99.9990 | 99.9982 | 99.9932 |
| y | 99.9897 | 99.9997 | 99.9593 | 99.9992 |
| z | 99.0673 | 99.7816 | 98.9156 | 96.4501 |
| X | 62.0358 | 88.0393 | 59.0551 | 19.7757 |
| δ_e | 68.6073 | 71.0806 | 66.5677 | 49.0182 |
| α | 73.6976 | 91.6101 | 73.2458 | 59.6657 |
| Average | 85.7905 | 93.8475 | 85.0293 | 74.1647 |

B-2-2 Sensitivity with regular gain $P_\theta = 0.4$ **Scenario 5**

99.9990458287532 99.9989286854248 99.9974142487824 99.9977246604377 99.9966564439905

Table B-6: Sensitivity of τ_1 with minimum θ_{ref} and normal gain P_θ

| | $\theta_{\text{ref}} = -3.5[\text{deg}]$ | | | |
|----------------|--|----------------|----------------|----------------|
| | -2JND τ_1 | -1JND τ_1 | +1JND τ_1 | +2JND τ_1 |
| q | 99.9967 | 99.9977 | 99.9990 | 99.9989 |
| u | 99.4000 | 99.5320 | 99.6326 | 99.5492 |
| v | 99.3734 | 99.5014 | 99.5816 | 99.4701 |
| w | 94.2816 | 96.9248 | 97.8707 | 95.0589 |
| θ | 98.3133 | 98.8652 | 98.5168 | 97.2379 |
| x | 99.9979 | 99.9982 | 99.9988 | 99.9991 |
| y | 99.9967 | 99.9977 | 99.9990 | 99.9989 |
| z | 99.7575 | 99.8108 | 99.7575 | 99.6211 |
| X | 93.3073 | 97.6540 | 98.8220 | 95.2501 |
| δ_e | 80.0370 | 88.6006 | 88.6778 | 79.2310 |
| α | 96.6705 | 98.3082 | 98.9046 | 97.2156 |
| Average | 96.4666 | 98.1083 | 98.3418 | 96.6028 |

Scenario 6**Table B-7:** Sensitivity of a_1 with minimum θ_{ref} and normal gain P_θ

| | $\theta_{\text{ref}} = -3.5[\text{deg}]$ | | | |
|----------------|--|-------------|-------------|-------------|
| | -2JND a_1 | -1JND a_1 | +1JND a_1 | +2JND a_1 |
| q | 99.9874 | 99.9971 | 99.9671 | 99.9975 |
| u | 95.9615 | 98.4920 | 98.0867 | 93.5821 |
| v | 95.8229 | 98.4647 | 97.6895 | 92.7225 |
| w | 44.9286 | 77.6760 | 57.7653 | 35.6936 |
| θ | 83.2882 | 94.0581 | 85.2968 | 67.6525 |
| x | 99.9921 | 99.9961 | 99.9993 | 99.9960 |
| y | 99.9874 | 99.9971 | 99.9671 | 99.9975 |
| z | 98.5617 | 99.4738 | 98.9117 | 96.4873 |
| X | 56.4971 | 83.5532 | 62.5415 | 23.1279 |
| δ_e | 60.9931 | 62.8415 | 65.2068 | 49.9068 |
| α | 66.0629 | 86.4835 | 74.4252 | 60.3356 |
| Average | 82.0010 | 91.0029 | 85.4445 | 74.4998 |

Scenario 7

Table B-8: Sensitivity of τ_1 with maximum θ_{ref} and normal gain P_θ

| | $\theta_{\text{ref}} = 1.5[\text{deg}]$ | | | |
|----------------|---|-----------------------|-----------------------|-----------------------|
| | -2JND τ_1 | -1JND τ_1 | +1JND τ_1 | +2JND τ_1 |
| q | 99.9995 | 99.9996 | 99.9978 | 99.9804 |
| u | 99.7610 | 99.8196 | 99.7417 | 99.5665 |
| v | 99.7269 | 99.7939 | 99.7252 | 99.5124 |
| w | 97.0492 | 98.8135 | 97.9106 | 94.5832 |
| θ | 98.4868 | 99.2240 | 99.3370 | 98.3791 |
| x | 99.9995 | 99.9995 | 99.9991 | 99.9987 |
| y | 99.9996 | 99.9996 | 99.9978 | 99.9804 |
| z | 99.8019 | 99.8831 | 99.8946 | 99.7901 |
| X | 95.5750 | 98.9361 | 97.8001 | 93.5079 |
| δ_e | 84.3116 | 91.4056 | 89.1281 | 79.2770 |
| α | 98.2887 | 99.3630 | 98.7628 | 96.6876 |
| Average | 97.5454 | 98.8397 | 98.3906 | 96.4802 |

Scenario 8

Table B-9: Sensitivity of a_1 with maximum θ_{ref} and normal gain P_θ

| | $\theta_{\text{ref}} = 1.5[\text{deg}]$ | | | |
|----------------|---|--------------------|--------------------|--------------------|
| | -2JND a_1 | -1JND a_1 | +1JND a_1 | +2JND a_1 |
| q | 99.9986 | 99.9986 | 99.9370 | 99.9870 |
| u | 98.0661 | 99.5409 | 97.1327 | 92.9987 |
| v | 97.9292 | 99.5335 | 96.7533 | 91.4803 |
| w | 56.9673 | 84.6053 | 58.4783 | 39.7298 |
| θ | 86.3863 | 95.4484 | 87.8060 | 73.7372 |
| x | 99.9990 | 99.9999 | 99.9946 | 99.9870 |
| y | 99.9986 | 99.9986 | 99.9370 | 99.9987 |
| z | 99.1023 | 99.7468 | 98.8849 | 96.5590 |
| X | 60.9565 | 86.6014 | 60.2898 | 21.0049 |
| δ_e | 67.5272 | 68.2848 | 64.9399 | 48.1476 |
| α | 73.4976 | 90.6153 | 73.7452 | 60.2655 |
| Average | 85.4936 | 93.1250 | 85.2687 | 74.8220 |

B-2-3 Sensitivity with maximum gain $P_{\theta} = 0.8$

Scenario 9

Table B-10: Sensitivity of τ_1 with minimum θ_{ref} and maximum gain P_{θ}

| | $\theta_{\text{ref}} = -3.5[\text{deg}]$ | | | |
|----------------|--|----------------|----------------|----------------|
| | -2JND τ_1 | -1JND τ_1 | +1JND τ_1 | +2JND τ_1 |
| q | 99.9924 | 99.9969 | 99.9974 | 99.9998 |
| u | 99.4532 | 99.6438 | 99.8617 | 99.8394 |
| v | 99.4392 | 99.6253 | 99.8365 | 99.7961 |
| w | 90.8368 | 94.6286 | 98.6386 | 97.0235 |
| θ | 97.5379 | 98.5176 | 99.3094 | 98.6146 |
| x | 99.9986 | 99.9989 | 99.9995 | 99.9997 |
| y | 99.9924 | 99.9969 | 99.9974 | 99.9998 |
| z | 99.7560 | 99.8560 | 99.9166 | 99.8446 |
| X | 90.1062 | 95.4420 | 99.7523 | 97.3749 |
| δ_e | 72.6817 | 83.7294 | 85.1570 | 73.3965 |
| α | 94.5154 | 96.8438 | 99.2713 | 98.2939 |
| Average | 94.9378 | 97.1165 | 98.3400 | 96.7439 |

Scenario 10

Table B-11: Sensitivity of a_1 with minimum θ_{ref} and maximum gain P_{θ}

| | $\theta_{\text{ref}} = -3.5[\text{deg}]$ | | | |
|----------------|--|-------------|-------------|-------------|
| | -2JND a_1 | -1JND a_1 | +1JND a_1 | +2JND a_1 |
| q | 99.9863 | 99.9924 | 99.9796 | 99.9974 |
| u | 95.9314 | 98.5522 | 98.4124 | 94.0230 |
| v | 95.7836 | 98.5320 | 98.0973 | 93.3011 |
| w | 36.3510 | 69.5224 | 60.4522 | 38.1879 |
| θ | 79.9488 | 91.9699 | 87.7364 | 71.3904 |
| x | 99.9934 | 99.9971 | 99.9992 | 99.9949 |
| y | 99.9863 | 99.9924 | 99.9796 | 99.9975 |
| z | 98.5053 | 99.4900 | 99.1762 | 96.9232 |
| X | 52.8176 | 79.7737 | 65.7659 | 26.2544 |
| δ_e | 51.1482 | 49.5456 | 59.1642 | 46.7987 |
| α | 60.9251 | 81.5122 | 75.8514 | 61.4626 |
| Average | 79.2168 | 88.0804 | 85.8758 | 75.3026 |

Scenario 11**Table B-12:** Sensitivity of τ_1 with maximum θ_{ref} and maximum gain P_θ

| | $\theta_{\text{ref}} = 1.5[\text{deg}]$ | | | |
|----------------|---|-----------------------|-----------------------|-----------------------|
| | -2JND τ_1 | -1JND τ_1 | +1JND τ_1 | +2JND τ_1 |
| q | 99.9947 | 99.9974 | 99.9926 | 99.9760 |
| u | 98.9822 | 98.9952 | 98.8102 | 98.5735 |
| v | 98.8494 | 98.8717 | 98.6714 | 98.3758 |
| w | 92.2510 | 94.5622 | 95.7827 | 93.5597 |
| θ | 95.6503 | 96.6750 | 97.7025 | 97.2678 |
| x | 99.9969 | 99.9965 | 99.9953 | 99.9946 |
| y | 99.9947 | 99.9974 | 99.9926 | 99.9760 |
| z | 99.4102 | 99.4984 | 99.5226 | 99.4183 |
| X | 92.8390 | 96.8794 | 97.7331 | 94.1009 |
| δ_e | 79.4750 | 88.4569 | 87.2284 | 75.6679 |
| α | 95.7506 | 97.1244 | 97.6554 | 96.1266 |
| Average | 95.7451 | 97.3685 | 97.5536 | 95.7323 |

Scenario 12**Table B-13:** Sensitivity of a_1 with maximum θ_{ref} and maximum gain P_θ

| | $\theta_{\text{ref}} = 1.5[\text{deg}]$ | | | |
|----------------|---|--------------------|--------------------|--------------------|
| | -2JND a_1 | -1JND a_1 | +1JND a_1 | +2JND a_1 |
| q | 99.9986 | 99.9922 | 99.9417 | 99.9844 |
| u | 98.1029 | 99.0682 | 95.6254 | 90.2485 |
| v | 97.8255 | 98.9513 | 95.0765 | 89.6086 |
| w | 54.9127 | 78.3669 | 58.9513 | 40.9167 |
| θ | 83.8036 | 92.2814 | 87.8192 | 75.6362 |
| x | 99.9992 | 99.9983 | 99.9879 | 99.9773 |
| y | 99.9986 | 99.9922 | 99.9417 | 99.9844 |
| z | 98.9289 | 99.4544 | 98.3972 | 96.0335 |
| X | 59.3296 | 83.1490 | 61.5186 | 22.0197 |
| δ_e | 63.8909 | 60.4519 | 59.7459 | 44.4273 |
| α | 72.6148 | 87.1893 | 73.4956 | 59.7008 |
| Average | 84.4915 | 90.8092 | 84.5952 | 74.4119 |

Part IV

Appendices to Final Report

Changes to Citation Stall Model

For the work in this thesis, there were several changes made to the Citation Model compared to the model used by Smets, De Visser, and Pool [9] and Imbrechts, De Visser, and Pool [10], which are described here.

C-1 Separating the Model from Simulink simulation

The original model as made by Smets, De Visser, and Pool [9] uses a single Simulink file to simulate both in MATLAB and in DUECA. In order to simulate in MATLAB, the trim output from the model should be fed back into the inputs, leading to a trimmed start of the stall maneuver. For DUECA implementation, this should be disconnected, as the trim is added to the simulation input in the CitationModel.cxx file in DUECA. To prevent any mishaps in connecting and disconnecting this trim input when preparing for any simulation, the model as needed for DUECA is separated from the trim feedback. As a result, the model as needed for DUECA is a standalone Simulink file.

In order to still be able to simulate the Citation in MATLAB as well, another Simulink file is created, which is named Citation_Simulink. Here, the trim output of the DUECA Citation model is connected to the input, such that trimmed flight is also established here without the DUECA interface. Now, the changes to the Citation model can be tested in MATLAB and without any changes compiled into C code for DUECA.

C-2 Separating roll autopilot from stall autopilot

In the original model, there was a single boolean to enable and disable the stall autopilot, which would switch off the input from the throttle, the pitch control which consists of the altitude hold mode and the recovery controller, and finally the roll control. However, this experiment still focuses on a symmetric stall and hence, the roll controller should still remain active in order to keep wings level throughout the maneuver. Therefore, an extra boolean was

added called `stall_ap_roll_enable` which can be separately switched off if needed. For this simulation, it is automatically set to true by DUECA, so the roll controller is always active, also during the active experiment.

C-3 Moved calculations for stall autopilot to the controller block

During the initial testing of the active experiment, an unexpected oscillatory behavior in pitch angle was found. This led to numerous corrections that were made to the original model as used before. Firstly, to allow a more rigorous debugging, the calculations necessary for the stall autopilot were removed from the Aerodynamic Model block. Initially, the calculation of when the recovery maneuver should start was performed in the Aerodynamic Model block, which would send a signal from the aerodynamic model block back to the controller block in which the stall autopilot is incorporated.

Now, the calculations for the stall autopilot are placed in the controller block, which leaves only the calculation of the flow separation point X in the aerodynamic block. There is still a mistake in the calculation of the flow separation point. The integrator that integrates \dot{X} to obtain X has an initial condition of 1, whereas simulation without the stall autopilot on show that the value of X rapidly decreases after starting the simulation, starting at approximately 0.8 instead of 1. This was deemed of little influence on the outcome of simulation for this specific research, but should be taken up in future work.

The changes can be seen in Figure C-1. The "determine start recovery" block is added, which came directly from the "Calculate flow separation point" block in the aerodynamic model block.

Next to this, the "Stall AP throttle setting" was changed. Originally, the boolean indicating full thrust was set in the "Recovery controller" block. This calculation was removed from the recovery controller and added to the throttle itself. This required a redesign of the throttle autopilot, which will be discussed next.

C-4 Matching throttle of stall autopilot and trim values

The original autopilot throttle block can be seen in Figure C-2. Here, the boolean $-T-$ comes from the recovery controller, which is now moved to the throttle block as discussed in the previous section, which can be seen in Figure C-3. There were additional changes made to the throttle block. Before, the throttle coming from the trim conditions was canceled out in this block, which was done by subtracting $ut0(1)$ from the signal. The minimum thrust setting ($utmin$) is always present. If full thrust is required by the stall autopilot, the boolean is multiplied with a constant $-K-$, which is the difference between the maximum thrust setting $utmax$ and $utmin$. This is then added to the minimum thrust setting, which results in a full thrust setting.

In this version, this correction of the trim thrust is taken out of the stall autopilot controller and simply corrected by the "deselect_thrust" gain in the controller loop (see Figure C-1). Now, the trim thrust, which is far too high to stall the aircraft, is not used. Still, the minimum thrust setting $utmin$ is taken as a baseline and the difference between $utmax$ and $utmin$ is

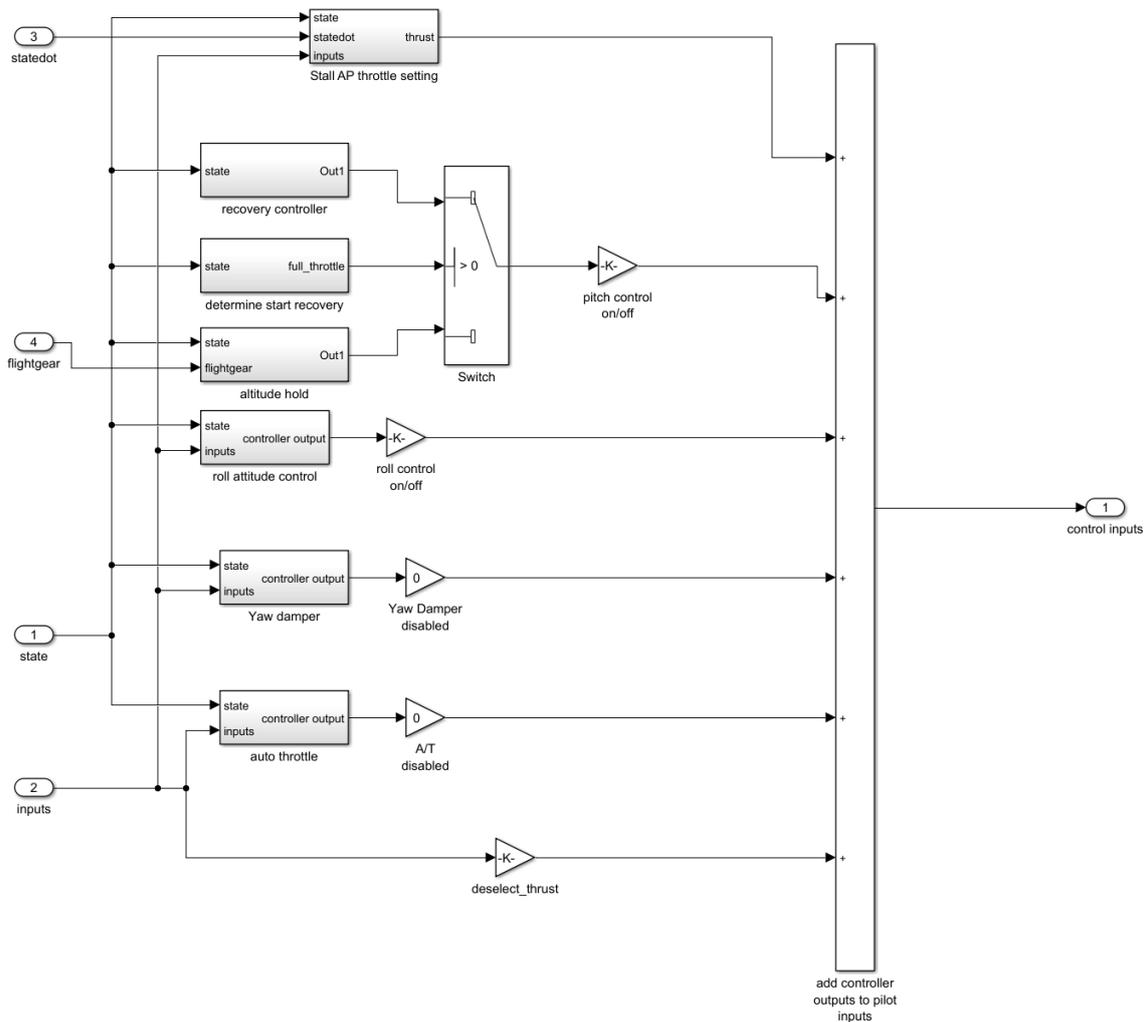


Figure C-1: The updated controller block with stall autopilot calculations included

added to this when full thrust is required. This can be seen in Figure C-3, where the $-K-$ is the same as in Figure C-2.

Furthermore, the pilot input is also multiplied by the same gain $-K-$. This requires the thrust from DUECA to be on the scale of 0 to 1 as well. This is taken care of in the DUECA software. If the throttle setting is below 0.5, the value is forced to zero. If the settings is above 0.5, the thrust setting is forced to 1. This is done to make the active simulation match the passive experiment in terms of thrust setting. Finally, another gain term can be seen, which switches the input from the autopilot and from the pedestal on the SIMONA Research Simulator based on the *stall_ap_enable* boolean.

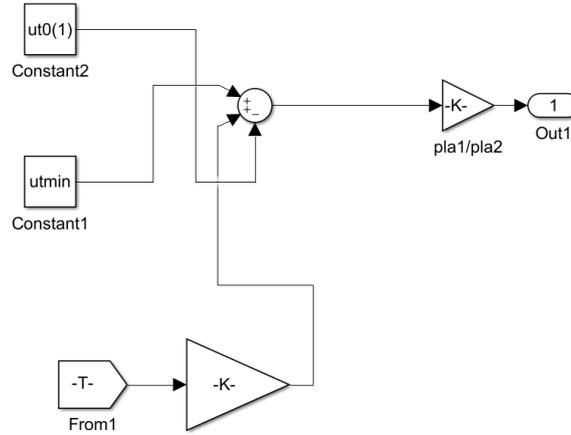


Figure C-2: Stall autopilot throttle settings as by Smets, De Visser, and Pool [9]

C-5 Correcting undamped short period

Upon further investigation into this oscillatory behavior found in the outcome of the active simulation, it was hypothesized that this could be due to the lack of a C_{m_q} term in the stall model as proposed by Van Ingen, De Visser, and Pool [25]. As discussed in the scientific paper, the symmetrical model structure is as follows:

$$C_L = C_{L_0} + C_{L_\alpha} \left(\frac{1 + \sqrt{X}}{2} \right)^2 \alpha + C_{L_{\alpha^2}} (\alpha - 6)_+^2 \quad (\text{C-1})$$

$$C_D = C_{D_0} + C_{D_\alpha} \alpha + C_{D_{\delta_e}} \delta_e + C_{D_X} (1 - X) + C_{D_{C_T}} C_T \quad (\text{C-2})$$

$$C_m = C_{m_0} + C_{m_\alpha} \alpha + C_{m_{X\delta_e}} \max\left(\frac{1}{2}, X\right) \delta_e + C_{m_{C_T}} C_T \quad (\text{C-3})$$

In the moment coefficient equation Equation (C-3), it can be seen that the pitch moment is dependent on the zero moment coefficient, the angle of attack, the elevator effectiveness, and the thrust. Without the stall autopilot active, this model behaves as shown in Figure C-4. As can be seen, after approximately 7.5 seconds, no additional oscillatory input is given in the elevator deflection, yet the pitch rate oscillates heavily.

Therefore, a C_{m_q} term is added to Equation (C-3). This C_{m_q} term will be multiplied by $\frac{qc}{V}$ and is based on the work by Van den Hoek, De Visser, and Pool [57]. They created a model for the Cessna Citation II for nominal flight conditions. Based upon this model, the interpolated value for C_{m_q} is obtained. This interpolation is done via Equation C-5, with the data as found in Table C-1. This interpolation based on $h = 5528m$ and $M = 0.2256$ resulted in an C_{m_q} of -8.1826 .

$$f(x, y) = \frac{1}{(x_2 - x_1)(y_2 - y_1)} \begin{bmatrix} x_2 - x & x - x_1 \end{bmatrix} \begin{bmatrix} f(x_1, y_1) & f(x_1, y_2) \\ f(x_2, y_1) & f(x_2, y_2) \end{bmatrix} \begin{bmatrix} y_2 - y \\ y - y_1 \end{bmatrix} \quad (\text{C-4})$$

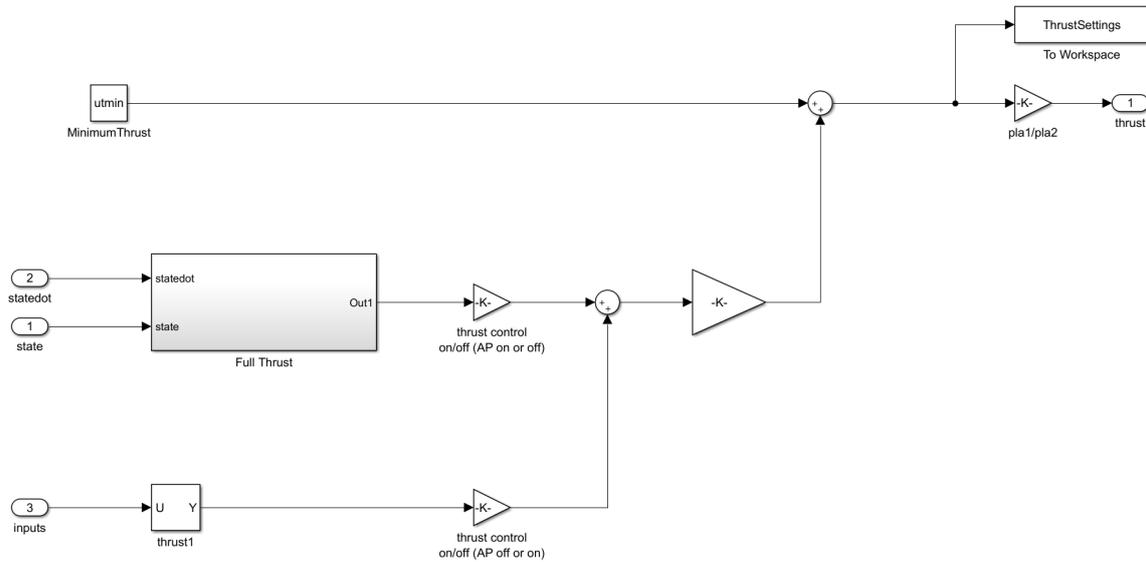


Figure C-3: Updated auto-throttle of the stall autopilot

Table C-1: Values for C_{m_q} as found by Van den Hoek, De Visser, and Pool [57] for different altitude and Mach number

| | $h = 5000\text{m}$ | $h = 6000\text{m}$ |
|---------|--------------------|--------------------|
| $M=0.2$ | -7.1095 | -7.1501 |
| $M=0.3$ | -11.2177 | -11.2583 |

The updated model was tested in the SIMONA Research Simulator with a pilot-it-the-loop evaluation. The pilot participating was a Cessna Citation II test pilot. This experiment gave the following pitch response as can be seen in Figure C-5. The results show that the pitch before the stall can be controlled without extreme control inputs, although the pitch does still slightly oscillate. However, after the stall, the oscillations are quite strong and are harder to control. The test pilot also mentioned that the oscillatory pitch response after the stall was very unnatural for the Cessna Citation II aircraft.

The fact that the oscillations are still present despite adding a C_{m_q} term is sensible, since the C_{m_q} is not fully integrated into the model. This requires a new selection of the right parameters for the stall model, thus repeating the work Van Ingen, De Visser, and Pool [25] have done with a C_{m_q} term forced into the model. This is deemed to be beyond the scope of this work.

However, this still leaves the problem of an uncontrollable pitch angle which could severely impact the ability of the participants to detect differences between a_1 as they will most likely shift their attention to keeping the pitch angle within limits. Therefore, it is opted to increase the C_{m_q} even further to make sure that the pitch angle is dampened.

To do this, the control inputs from the test pilot were isolated from the SIMONA data and used as inputs for offline Simulink simulations. The inputs of all 10 test runs with the test pilot were used for these simulations. Several higher values for C_{m_q} were tested, as can be

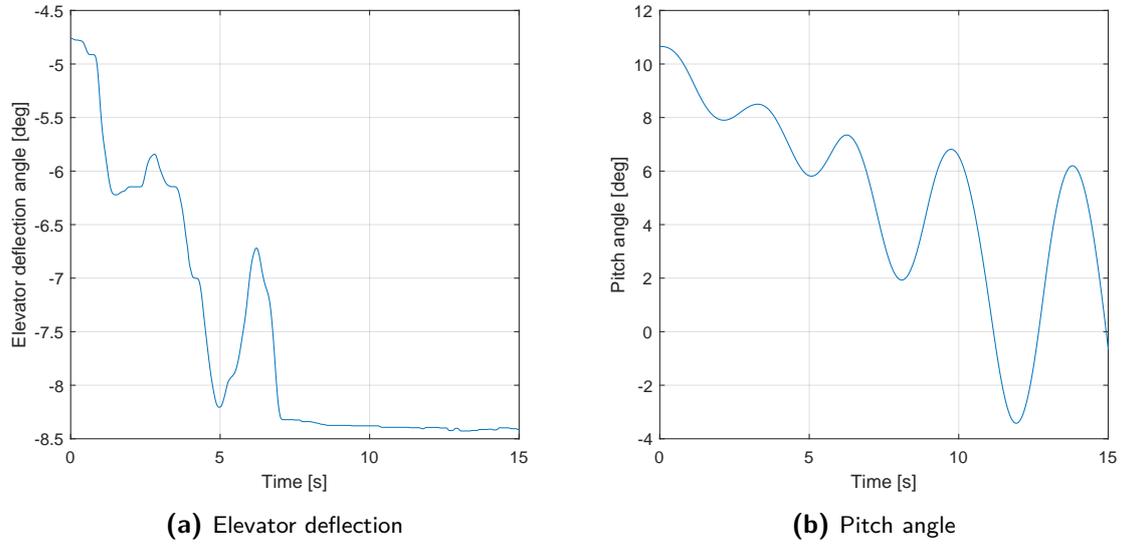


Figure C-4: Test run with no C_{m_q} with control input

seen in Figure C-6. Based upon the results of the simulations with the various C_{m_q} settings, a final setting of $C_{M_q} = -22$ was chosen, as this value gave the least oscillations without over-damping.

Finally, as a verification, the updated model with $C_{M_q} = -22$ was tested in a pilot-in-the-loop evaluation with the same test pilot as in the previous experiment. He mentioned that the handling qualities of the updated model were much better and that they felt more natural than the previous model. As a result, these updated settings will be used in the experiment.

As a final verification, the eigenvalues of the pitch-elevator input transfer function were analyzed. To do this, the system without trim loop was linearized in MATLAB through the *linmod* function to obtain a state space representation. From this, the correct input and output blocks were selected to retrieve the state space system with only q as output and δ_e as input. The redundant states were removed from the analysis through the *minreal* function, which obtained the minimal realization through pole-zero cancellation. A tolerance was set to $1 \cdot 10^{-6}$. It was verified that there was no significant information lost in the analysis of the reduced system, by analyzing the impulse response of both systems. This can be seen in Figure C-7.

With this verified, the eigenvalues were analyzed. In the system without C_{m_q} , a positive pair of eigenvalues was found: $2.42 \cdot 10^{-3} \pm 0.192i$. From the flight dynamics reader[58], it was found that the Cessna Citation II should have a damped short period with eigenvalues of $-3.9161 \cdot 10^{-2} \pm 3.7971 \cdot 10^{-2}i$ in nominal flight conditions. The updated model with $C_{m_q} = -22$ has short period eigenvalues of $-7.19 \cdot 10^{-3} \pm 0.142i$. Although the magnitude of these eigenvalues is half of those presented in the flight dynamics reader, this was deemed acceptable for the current experiment.

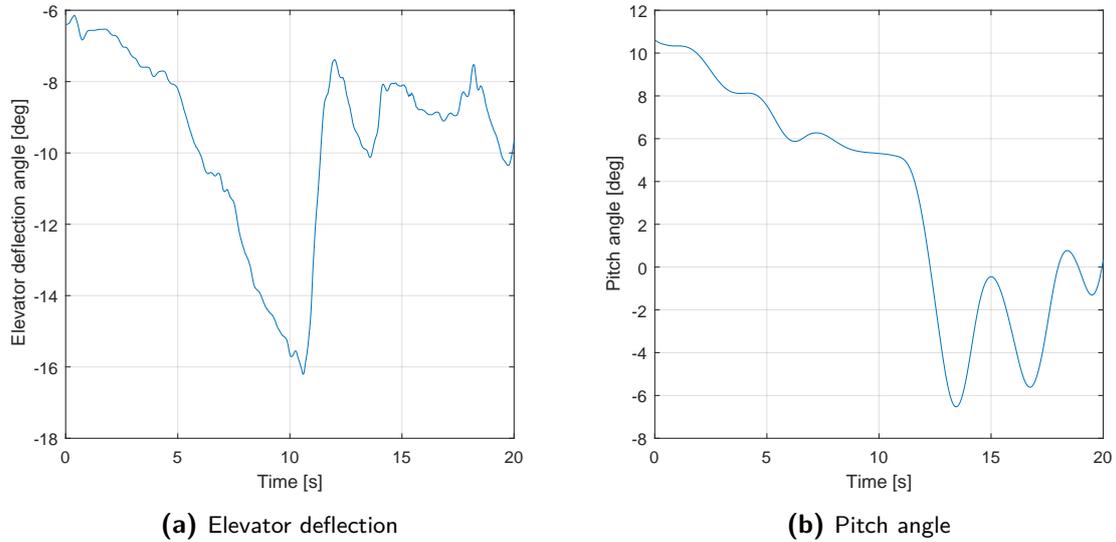


Figure C-5: Test run with $C_{M_q} = -8.1826$ with control input

C-6 Setting the Initial Conditions

Parallel to the work regarding the C_{m_q} , an additional effort was put into setting the initial conditions right. It was noticed that even with an added C_{m_q} , the model still showed oscillations at the start of the experiment, albeit that the oscillations were indeed damped over time. This can be seen in Figure C-8. It was hypothesized that this was due to the initial conditions and the fact that the model might not have been fully trimmed for this specific flight condition.

Consequently, the initial pitch rate q and elevator deflection δ_e were trimmed for each flight condition in the active experiment. This was done manually by adjusting the value for both the initial pitch rate and an additional elevator trim such that the oscillations as seen in Figure C-8b were minimized. This iterative process resulted in the pitch angle for each of the active conditions as seen in Figure C-9. The values for q_0 and $\delta_{e_{\text{extra}}}$ can be seen in Table C-2. This was only done for the active experiment, as the autopilot was able to keep the pitch attitude at the desired angle without oscillations.

Table C-2: Additional trim values for q_0 and $\delta_{e_{\text{extra}}}$ for active experiment

| a_1 | q_0 | $\delta_{e_{\text{extra}}}$ |
|---------|--------|-----------------------------|
| 27.6711 | -0.021 | -0.0105 |
| 28.7427 | -0.021 | -0.0105 |
| 29.8143 | -0.020 | -0.01025 |
| 30.8858 | -0.019 | -0.01025 |
| 31.9574 | -0.018 | -0.01 |
| 34.1005 | -0.016 | -0.009 |
| 36.2437 | -0.015 | -0.0075 |
| 38.3869 | -0.014 | -0.007 |

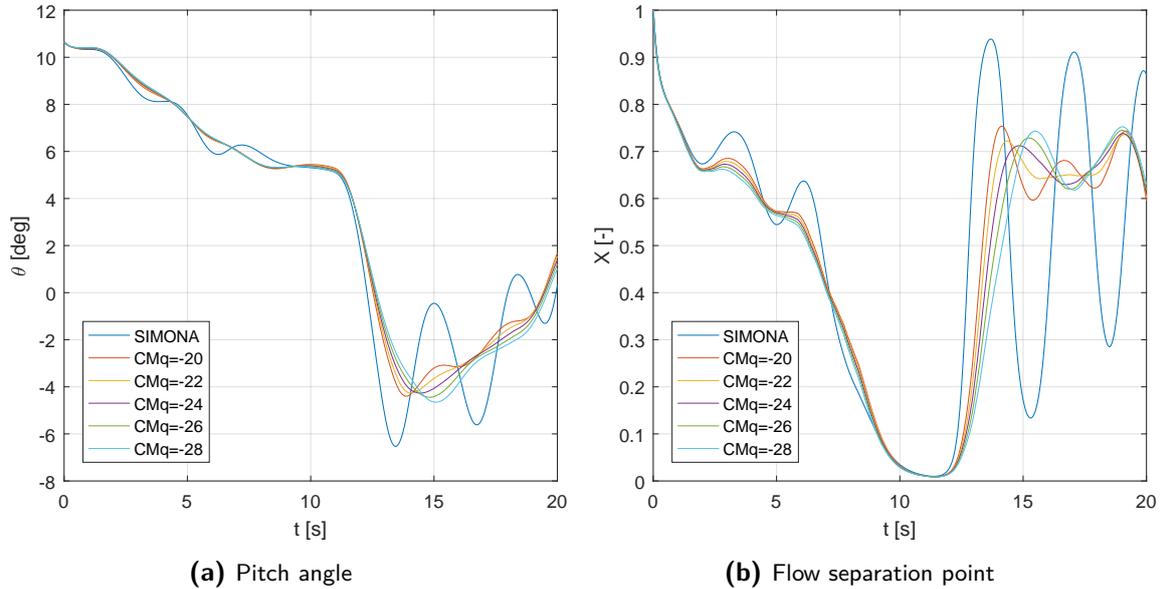


Figure C-6: Test run with $C_{Mq} = -20 \rightarrow -28$ with control input as shown in Figure C-5a

C-7 Analysis of the Updated Model

Finally, the model by Smets, De Visser, and Pool [9] with the baseline a_1 setting and the JND a_1 setting of 31.9574 were compared to the results of the same a_1 conditions for this experiment, to ensure that there were not specific states that had higher variations compared to Smets, De Visser, and Pool [9], as this would prevent a direct comparison of the thresholds. Therefore, several of the key states were compared.

It is important to note that what is analyzed here is if the difference between the baseline model with $a_1 = 27.6711$ and the JND value of Smets, De Visser, and Pool [9] with $a_1 = 31.9574$ is similar to these differences in the model used by Smets, De Visser, and Pool [9]. As visible in Figure C-10b, the difference minimum pitch angle during the recovery, as well as the difference between pitch angle before the recovery was initiated, is similar. Similar trends can be seen in the pitch rate, and flow separation point, shown in Figure C-12a and C-13a respectively.

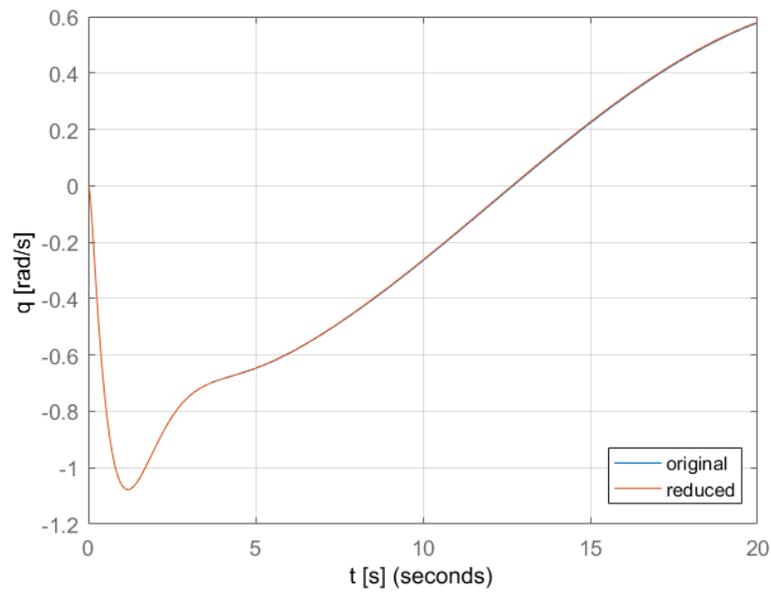
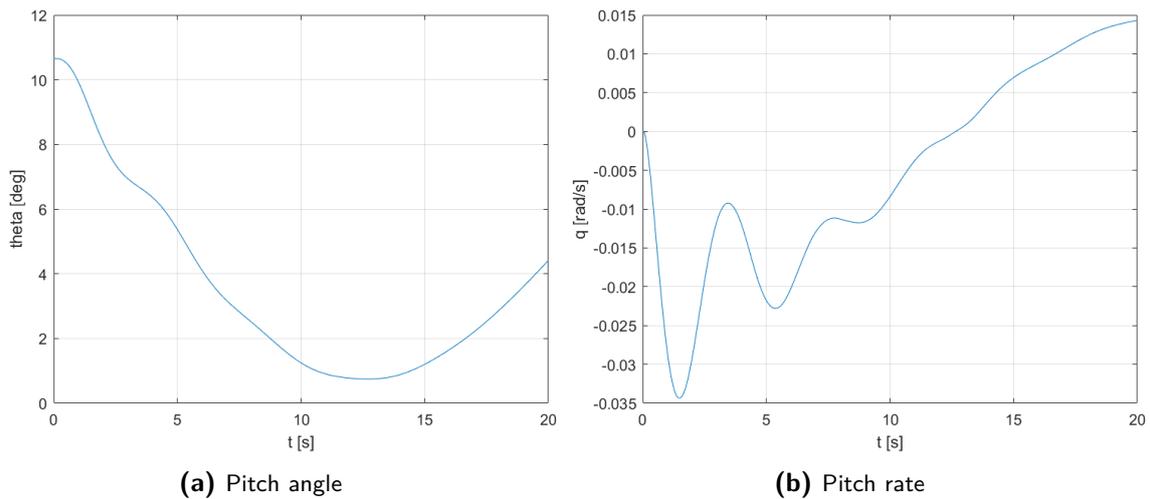


Figure C-7: Impulse response of the original $q - \delta_e$ system and the reduced system



(a) Pitch angle

(b) Pitch rate

Figure C-8: Simulation run with no inputs

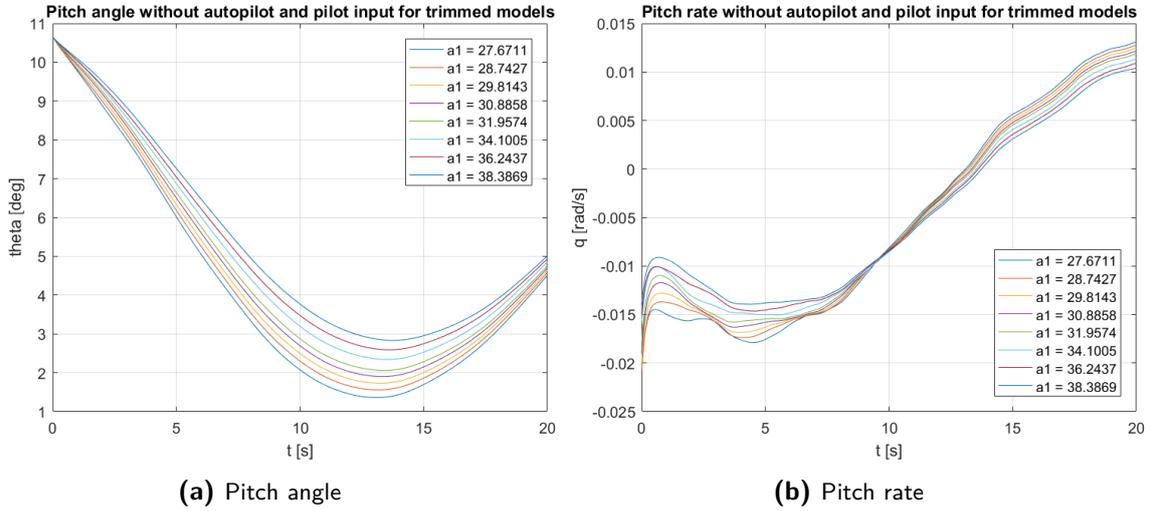


Figure C-9: Simulation runs with all active values for a_1 with trimmed pitch rate q and elevator deflection δ_e

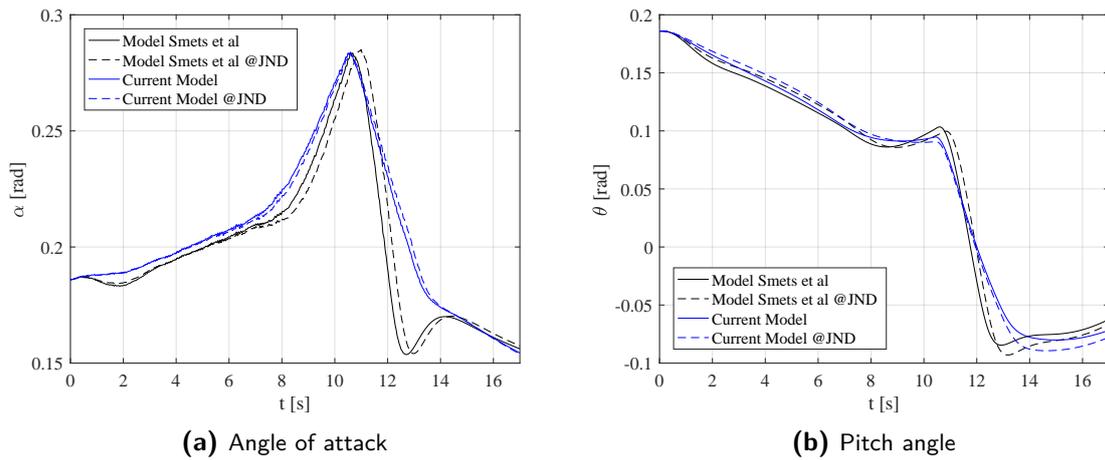


Figure C-10: Comparison of model by Smets et al [9] and the model used in this research for angle of attack and pitch angle

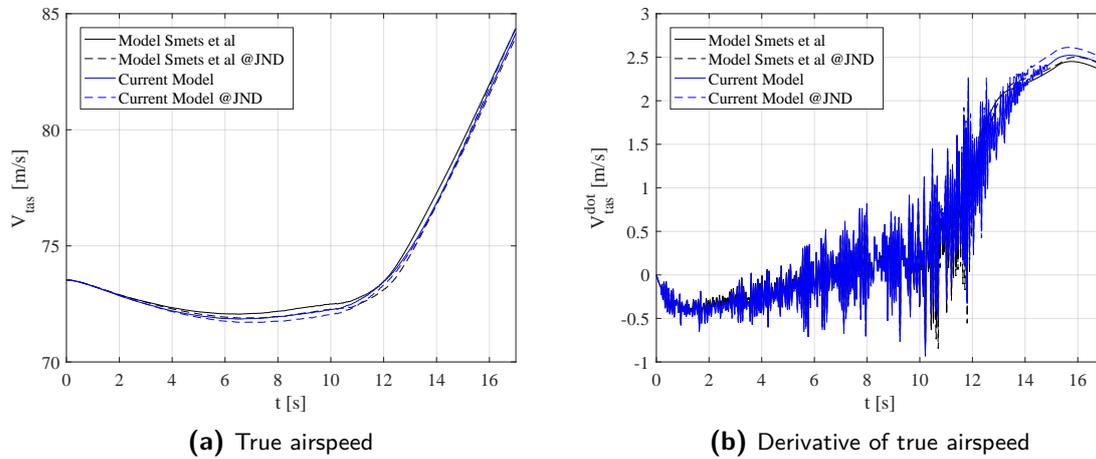


Figure C-11: Comparison of model by Smets et al [9] and the model used in this research for true airspeed and acceleration

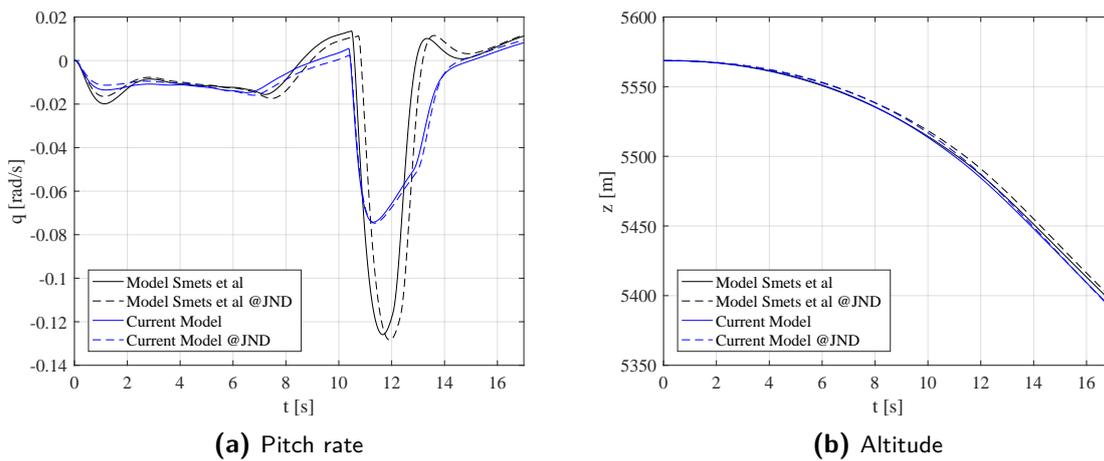


Figure C-12: Comparison of model by Smets et al [9] and the model used in this research for pitch rate and altitude

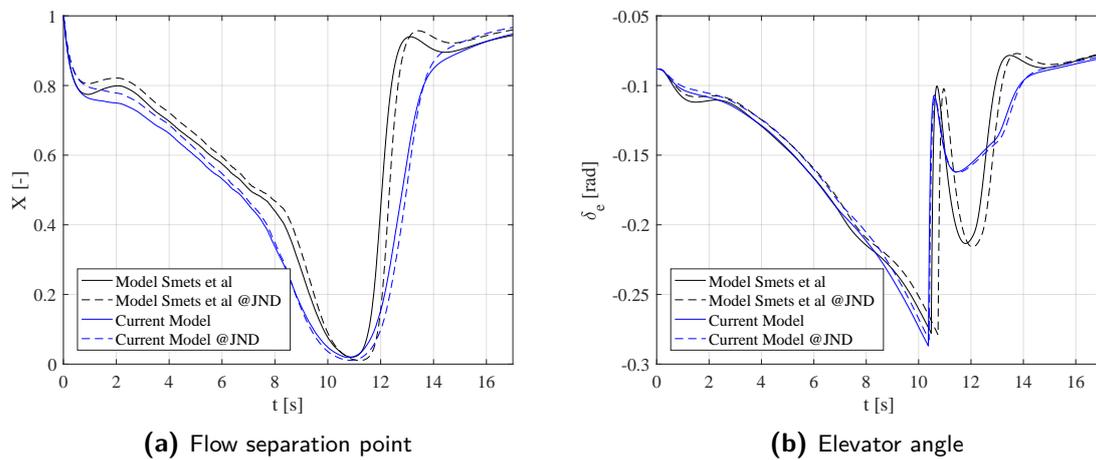


Figure C-13: Comparison of model by Smets et al [9] and the model used in this research for flow separation point and elevator angle

Appendix D

DUECA implementation of the experiment

D-1 The updated staircase procedure and ECI

The staircase procedure and Experiment Control Interface (ECI) for this experiment is updated with regards to the work by Smets, De Visser, and Pool [9] and Imbrechts, De Visser, and Pool [10]. In their work, the ECI handled all the processes, essentially being the "brain" of the program.

For this research, a different approach was taken. The ECI is now only the user interface between the ECI screen (see Figure D-1) and DUECA, setting all parameters for the staircase, as well as retrieving the sequence file for the active experiment. It sends most of its information to the new Staircase module, which is taken as inspiration from Vuik [59].

This Staircase module is the new "brain" of the program, handling all logic and experiment conditions. It uses the information from the ECI to set all conditions right and consequently executes the program based on inputs by the control column in SIMONA, as explained in the research paper. Furthermore, during offline simulations, the flexistick interface can be used for this. The important data streams between modules can be seen in Figure D-2

The Staircase module is programmed to handle a two-up, one-down procedure, but this can easily be transformed to another type of staircase, such as three-up, one-down, or a simple one-up, one-down.

D-2 How to use the ECI

The entire experiment can be controlled from the ECI, which is shown in Figure D-1, when *dusime* is in Hold. The sliders on top can be used to control different modes of the experiment. The top slider controls whether the experiment will be run in active or passive mode. By

Main settings

Subject number:

Passive Experiment Active Training

SIMONA settings

SIMONA Motion:
 Motion ON
 Motion OFF

Training Options:
 Single Comparison

Autopilot and stall model settings

Autopilot on or off:
 On
 Off

| Stall model settings | | # trial | label |
|----------------------|-----------|----------------|-------|
| tau1 | 0.254700 | # reversals | label |
| tau2 | 0.017600 | # correct 2U1D | label |
| a1 | 27.671100 | # correct | label |
| alpha | 0.208400 | Curr stepsize | label |
| Correct ans | label | Stall AP | label |
| Actual ans | label | | |

Choose file for active experiment:

Staircase Settings

| | | | | | | | |
|------------------|-----------------------------------|----------------------------------|----------------------------------|----------------|--------------------------------|----------------------------------|----------------------------------|
| Start value a1 | <input type="text" value="50.0"/> | <input type="button" value="-"/> | <input type="button" value="+"/> | # reversals | <input type="text" value="0"/> | <input type="button" value="-"/> | <input type="button" value="+"/> |
| Start stepsize | <input type="text" value="7.50"/> | <input type="button" value="-"/> | <input type="button" value="+"/> | # correct 2U1D | <input type="text" value="0"/> | <input type="button" value="-"/> | <input type="button" value="+"/> |
| # trials | <input type="text" value="0"/> | <input type="button" value="-"/> | <input type="button" value="+"/> | # correct | <input type="text" value="0"/> | <input type="button" value="-"/> | <input type="button" value="+"/> |
| Simulation Time: | <input type="text" value="17.0"/> | <input type="button" value="-"/> | <input type="button" value="+"/> | | | | |

Experiment control

EXPERIMENT START

Figure D-1: ECI for this experiment

sliding this, the Autopilot on or off button will automatically change as well. However, in case the active scenario needs to be tested with the autopilot on, this can be manually changed.

The slider below that is used to activate the "Training Options" part of the ECI, simultaneously hiding the "Apply" button below, as this button controls the experiment conditions and not the training. The training run can have a single test run on baseline conditions, or a comparison with the baseline in the first run and the maximum a_1 for the second run. By pushing "Apply Train", the experiment conditions are sent to the Staircase module, which sets everything for the simulation. Then, the "AP disconnect" button on the control column, or the "Start Train" button on the ECI can be used to start the simulation. After the training run has been completed and another training run is desired, the "Apply Train" button must be pressed again before the simulation can be started again. This repetitive step is only necessary during the training phase. When the training phase is finished, the slider can be used to go back to experiment conditions. This should reactivate the "Apply" button below as well. If this is not the case, slide the "Passive-Active" slider back and forth and this should work.

If a passive experiment will be started, the "Staircase Settings" can be used to set the initial settings for the staircase. The number of trials, reversals, correct 2UID, and correct answers can be set as well, to allow for a restart of the experiment at a specific point, if necessary. During the experiment itself, the "Stall model settings" can be used to monitor the experiment. All "label" entries will be updated during the simulation to reflect the progress the participant has made.

For an active experiment, first a file must be chosen in the "Choose file for active experiment:" button. This file should be a .txt file with all the required a_1 settings for the active experiment. The first entry of the file should be the length of the file, followed by all a_1 settings for each comparison. When a file is chosen, the "Apply" button becomes active again.

By pressing the "Apply" button, all settings are sent to the Staircase module, which, in turn, processes all information and distributes all information over the simulation. When adjustments are needed to the simulation, or a switch from passive to active must be made, the "Apply" button can be pushed again (only in Hold!) and the new information will be processed and sent.

D-3 SIMULINK Citation Model

The SIMULINK model of the stall model was altered as well. The model is a standalone SIMULINK file that can be exported to C code. Furthermore, another SIMULINK environment is created that imports the model and sets it up for offline simulations in Matlab itself. Consequently, there is no need anymore to switch on and off certain parts of the simulation when exporting it to C code.

If there remain any doubts after reading this documentation and comments added to the DUECA code, feel free to reach out to the author to ask any questions.

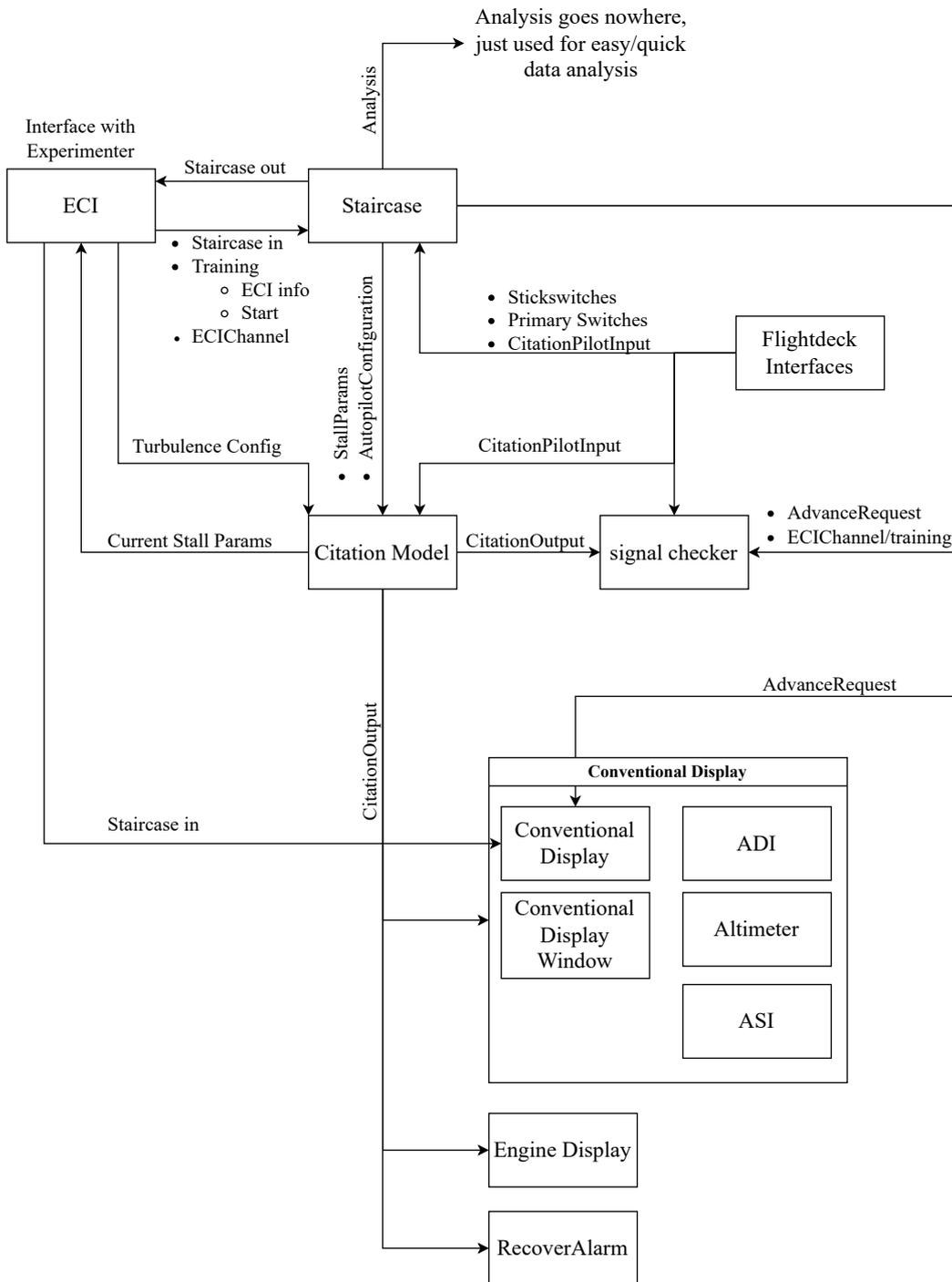


Figure D-2: Flowdiagram of the DUECA modules in the Citation Stall folder

Appendix E

Experiment Document

This chapter contains the documents provided to the participants for their experiments. It includes:

1. Experiment Consent form
2. Participant briefing
3. Questionnaire

Experiment Consent Form

Just Noticeable Differences for Quasi-Steady Stall Models in Active Flying

I hereby confirm, by ticking each box, that:

1. I volunteer to participate in the experiment conducted by the researcher (**Sybre Bootsma**) under supervision of **dr.ir. Daan Pool** from the Faculty of Aerospace Engineering of TU Delft. I understand that my participation in this experiment is voluntary and that I may withdraw from the study at any time, for any reason.
2. I have read the experiment briefing and confirm that I understand the experiment instructions and have had all remaining questions answered to my satisfaction.
3. I understand that taking part in the experiment involves two different test conditions. In one condition, I will be performing an observation task in the SIMONA Research Simulator. Only the simulation settings and the answers that I give regarding the noticeability of a certain difference between runs are saved.
4. I understand that taking part in the experiment involves two different test conditions. In the other condition, I will be performing an active control task in the SIMONA Research Simulator. The simulation settings, the answers that I give regarding the noticeability of a certain difference between runs, as well as the control inputs are saved.
5. I confirm that the researcher has provided me with detailed safety and operational instructions for the SIMONA Research Simulator (simulator setup, flight instrumentation, flight controls, fire escape ladder) used in the experiment. Furthermore, I confirm that I have understood the researcher's instructions for guaranteeing that the experiment will be performed in line with current RIVM COVID-19 guidelines and that this experiment shall always follow these RIVM guidelines.
6. I understand that the researcher will not identify me by name in any reports or publications that will result from this experiment, and that my confidentiality as a participant in this study will remain secure. Specifically, I understand that any demographic information I provide (age, pilot license type, flight hours, etc.) will only be used for reference and always presented in aggregate form in scientific publications. Furthermore, I understand that all subjective and objective measurement data will be stored under an anonymized participant number. Finally, I understand that I have the right to ask to have all of my data removed within 4 weeks of concluding the experiment.
7. I understand that this research will not, in any way, evaluate my performance regarding stall recovery procedures. The recovery data will solely be used to verify that the felt difference was due to differences in the model.

Contact information researcher:
Sybre Bootsma

Contact information research supervisor:
dr. ir. Daan Pool

8. I understand that the anonymised data may be used for future research carried out at the TU Delft.
9. I understand that this research study has been reviewed and approved by the TU Delft Human Research Ethics Committee (HREC). To report any problems regarding my participation in the experiment, I know I can contact the researchers using the contact information below.

My Signature

Date

My Printed Name

Signature of researcher

Contact information researcher:
Sybren Bootsma

Contact information research supervisor:
dr. ir. Daan Pool

Participants briefing:

Just Noticeable Differences for Quasi-Steady Stall Models in Active Flying

Within the Control & Simulation section at the Faculty of Aerospace Engineering, a stall task force group aims to develop an accurate model of the Cessna Citation II which can be used for effective stall recovery training. With this human-in-the-loop experiment we want to get a better understanding of the required accuracy of the critical stall model parameters and how active control influences this required accuracy.

The experiment will be performed in the SIMONA Research Simulator (SRS) at the Faculty of Aerospace Engineering at TU Delft, see Figure 1. You, as a participant, will be seated in the left hand seat of the simulator and be provided with motion, an outside visual environment representation, an instrument panel where you can read airspeed, altitude, vertical speed, heading, etc. and an additional display with the engine settings. Figure 1 shows the pilot station and simulator cabin. You will wear a noise cancellation headset to cancel out any noise coming from the actuators when moving you around in space.

In this experiment we will investigate two different cases. For both cases, we follow a procedure where we ask you to detect which of the two consecutive simulated stall maneuvers has the most abrupt stall.

In one case, you will only be an observer. Here, a pre-programmed autopilot will fly the aircraft into a stall and perform the recovery procedure. You will be a "passenger" in this case. In this way, the stall is always performed in the same way, allowing you to fully focus on detecting which stall simulation run had a more abrupt stall. We will refer to this case as the "Passive Experiment".

In the other case, you will have control over the aircraft. This will be referred to as the "Active Experiment". During the active experiment, you will use the SIMONA simulator's flight controls to follow a reference bar on the primary flight display, indicating the desired pitch attitude such that the aircraft will stall. Once the correct stall angle is reached, a message "RECOVER" will appear on the primary flight display, as well as a audio message saying "Recover", after which you can recover the aircraft and return to the nominal flight path once recovered. The simulation will end automatically.

The procedure of each case will be as follows:

1. You will experience two consecutive simulated stalls (each one takes approx. 15 seconds)
2. You push either up or down on the control column, indicating you think that the first or the second stall was more abrupt. You may, at some point, be unsure which of the two runs was more abrupt. In this case, you still need to make a decision which you think was more abrupt.
3. Once you have made your decision, you will press the start button, which will restart the procedure as described by all of the points above. If you made a decision quickly and pushed the start button quickly as well, the start may be delayed a bit. This is normal, as the simulator needs some time to reset. When the simulator is ready, it will start automatically.
4. The researcher will define when sufficient data is collected and will consequently stop the procedure.

Before each of the cases starts, we will do some familiarization runs, also named "Training Runs". For both experiments, we will explain and train you until the experiment is clear and you are confident in detecting differences in the abruptness of the stall. After this, we will start the measurements. The researcher will indicate if you start with the active experiment followed by the passive experiment, or vice versa.

The structure of the experiment will be as follows:

1. Training runs (+/- 15 min.)
2. Active or Passive experiment (+/- 45 min.)
3. Break (+/- 15 min.)
4. Training runs (+/- 15 min.)
5. Passive or Active experiment (+/- 45 min.)
6. End of the experiment + debriefing

The entire experiment will take around 2.5 hours, including the breaks. Note that your participation in the experiment is completely voluntary.



(a) Outside of SIMONA



(b) Inside of SIMONA

Figure 1: SIMONA research simulator

Demographic Information for Active JND Stall experiment

1. Participant Number _____ (filled in by the experimenter)
2. Age _____
3. Gender _____
4. Current Type rating _____
5. Total amount of flight hours _____
6. Years of flying experience _____
7. Current role: Captain / First Officer / Second Officer / Other _____ (please specify)
8. Are you/Have you been a flight instructor? Yes / No
9. Are you/Have you been a test pilot? Yes / No

Please note that this information will not be linked to you, and it will only be presented in the Scientific work as aggregated data.

Appendix F

Individual results

Subject 1

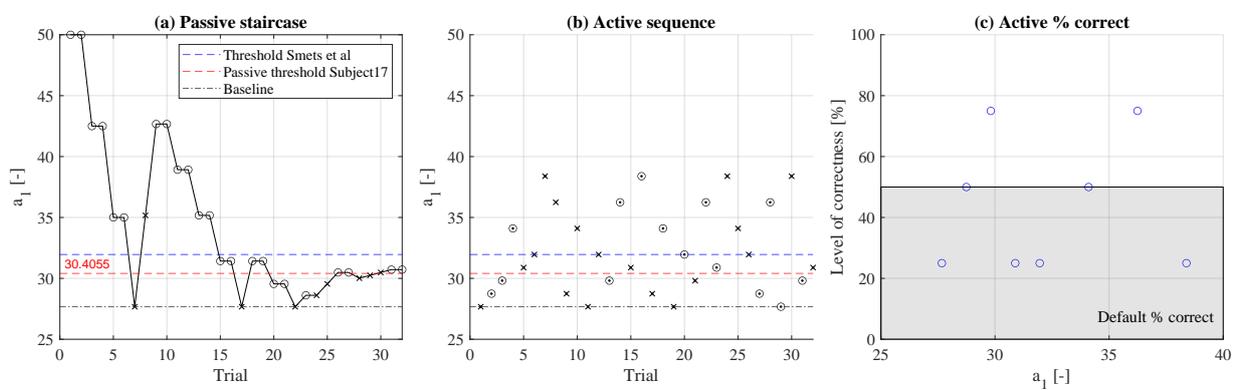


Figure F-1: Results of both the passive and active experiment for Subject 1, presented together with the percentages correct for the active experiment.

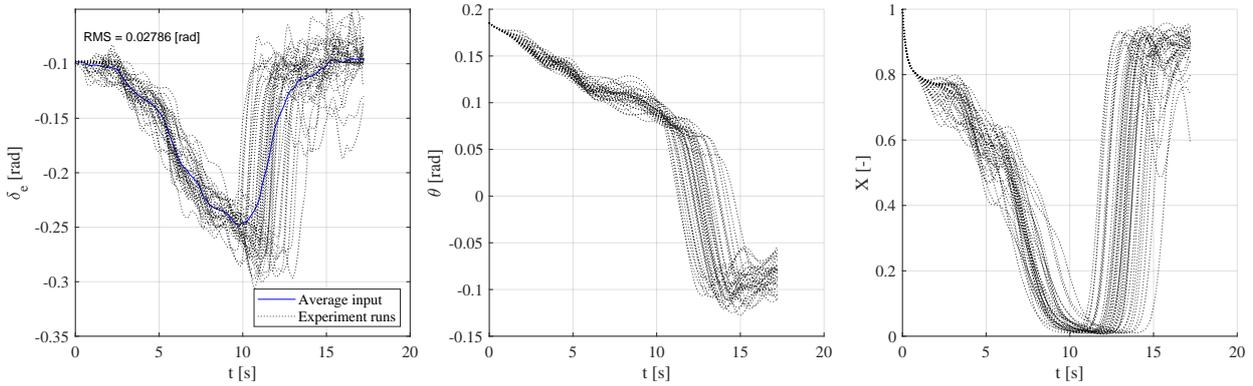


Figure F-2: Elevator inputs for all baseline runs from Subject 1, together with the pitch angle and flow separation point.

Subject 2

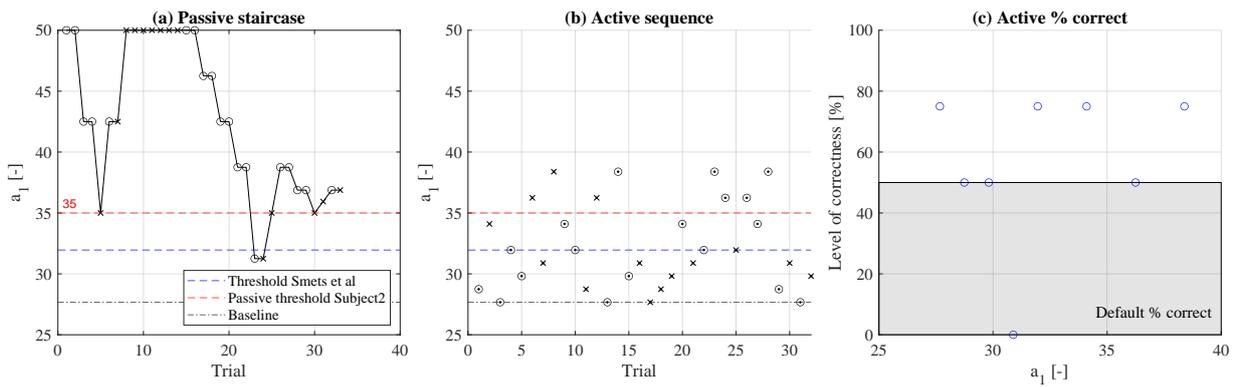


Figure F-3: Results of both the passive and active experiment for Subject 2, presented together with the percentages correct for the active experiment.

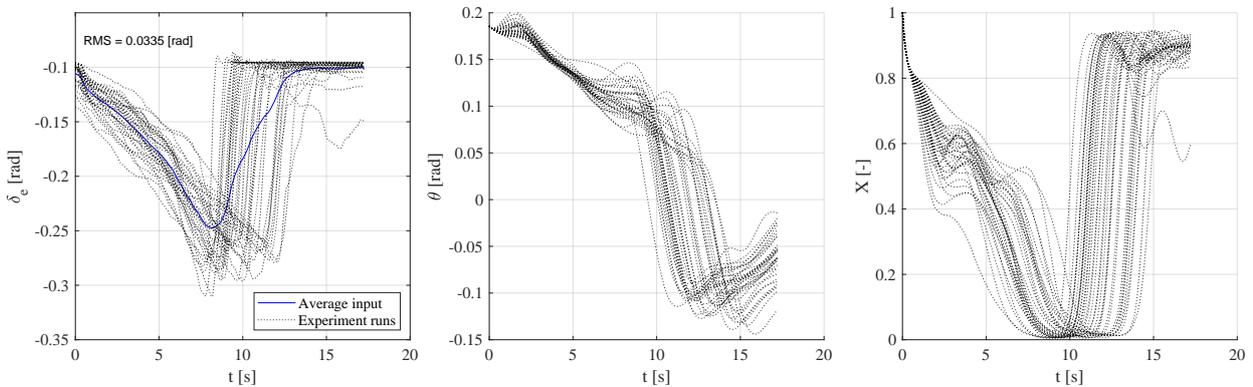


Figure F-4: Elevator inputs for all baseline runs from Subject 2, together with the pitch angle and flow separation point.

Subject 3

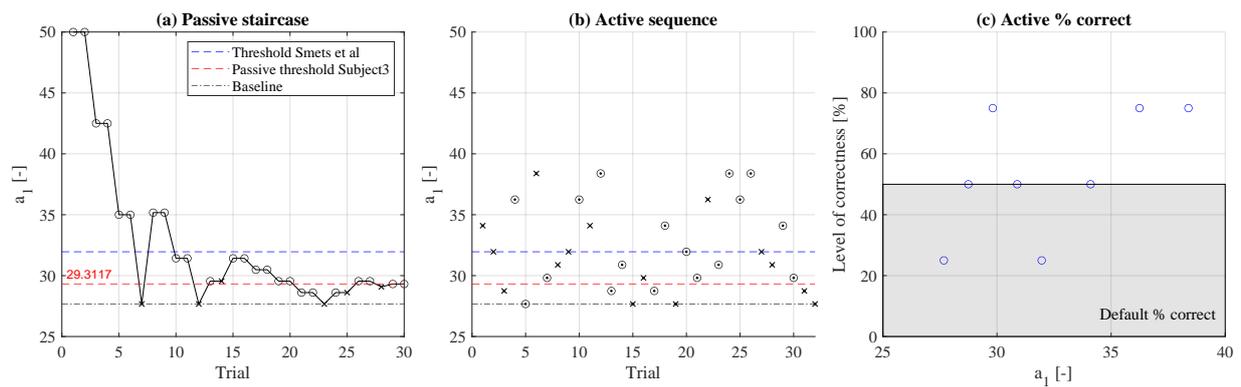


Figure F-5: Results of both the passive and active experiment for Subject 3, presented together with the percentages correct for the active experiment.

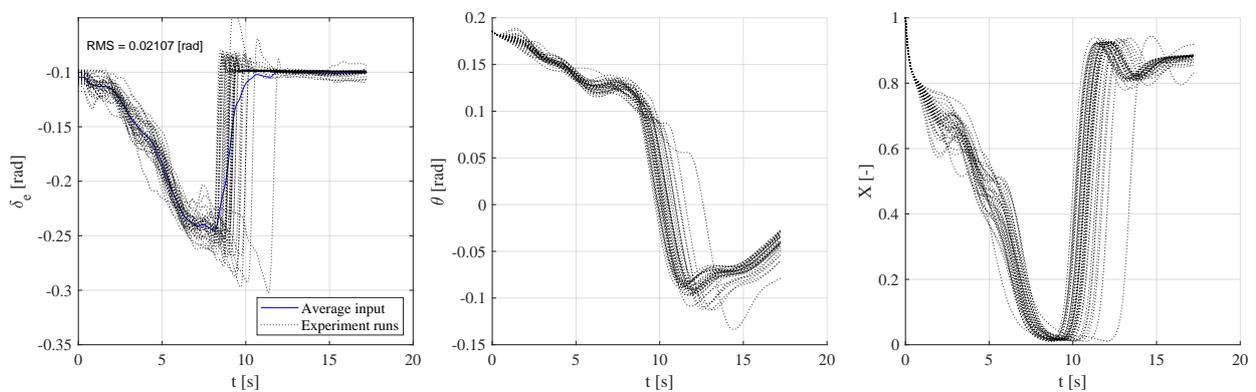


Figure F-6: Elevator inputs for all baseline runs from Subject 3, together with the pitch angle and flow separation point.

Subject 4

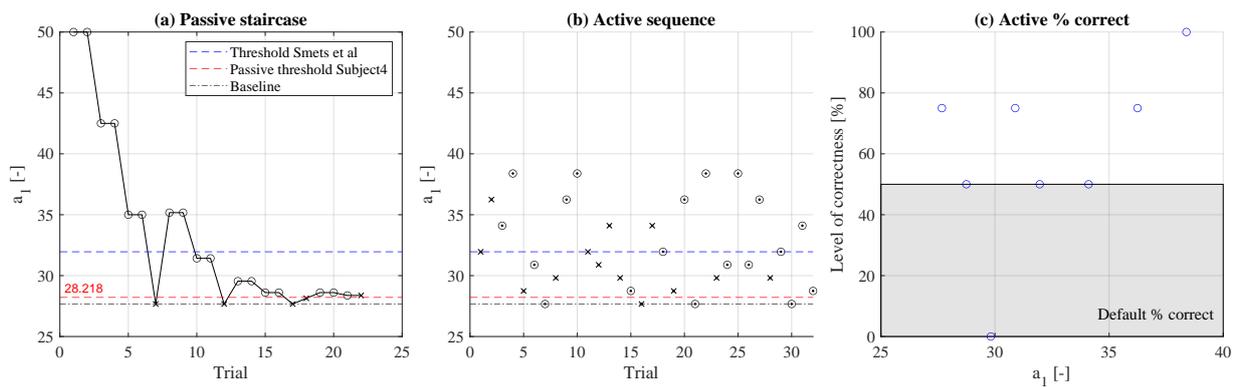


Figure F-7: Results of both the passive and active experiment for Subject 4, presented together with the percentages correct for the active experiment.

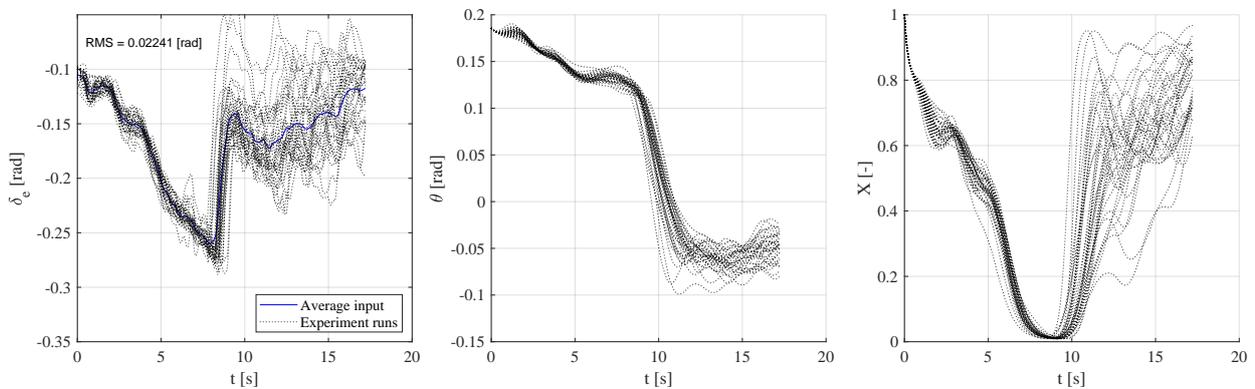


Figure F-8: Elevator inputs for all baseline runs from Subject 4, together with the pitch angle and flow separation point.

Subject 5

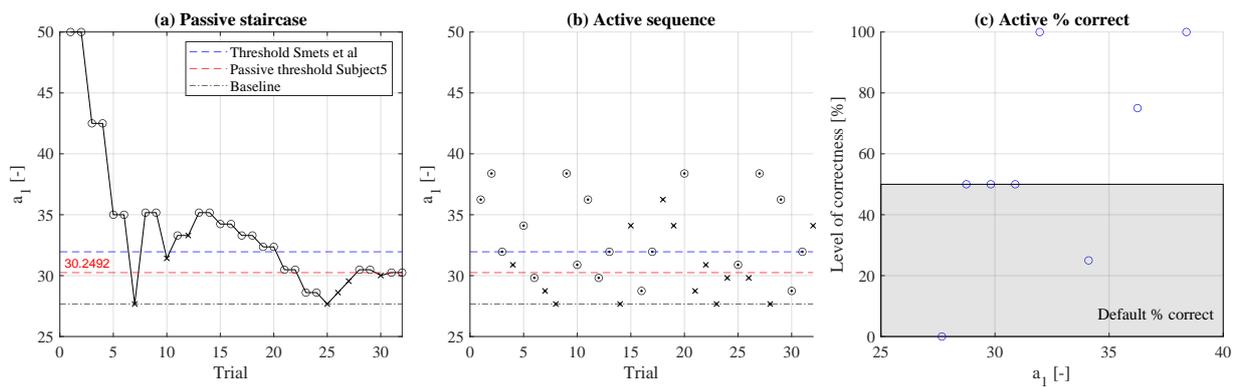


Figure F-9: Results of both the passive and active experiment for Subject 5, presented together with the percentages correct for the active experiment.

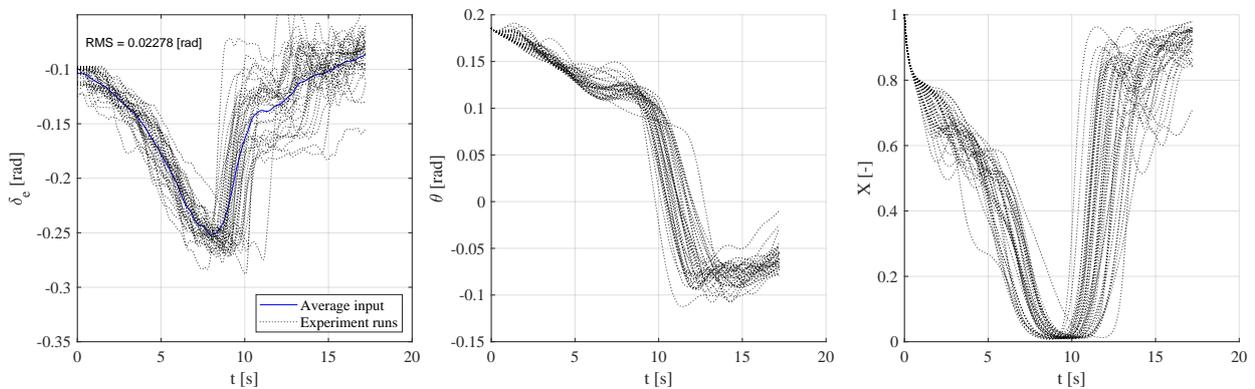


Figure F-10: Elevator inputs for all baseline runs from Subject 5, together with the pitch angle and flow separation point.

Subject 6

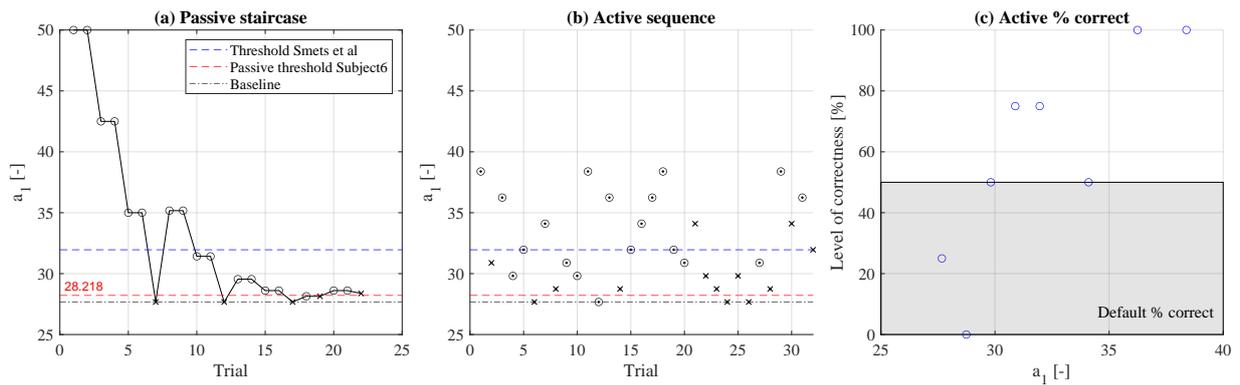


Figure F-11: Results of both the passive and active experiment for Subject 6, presented together with the percentages correct for the active experiment.

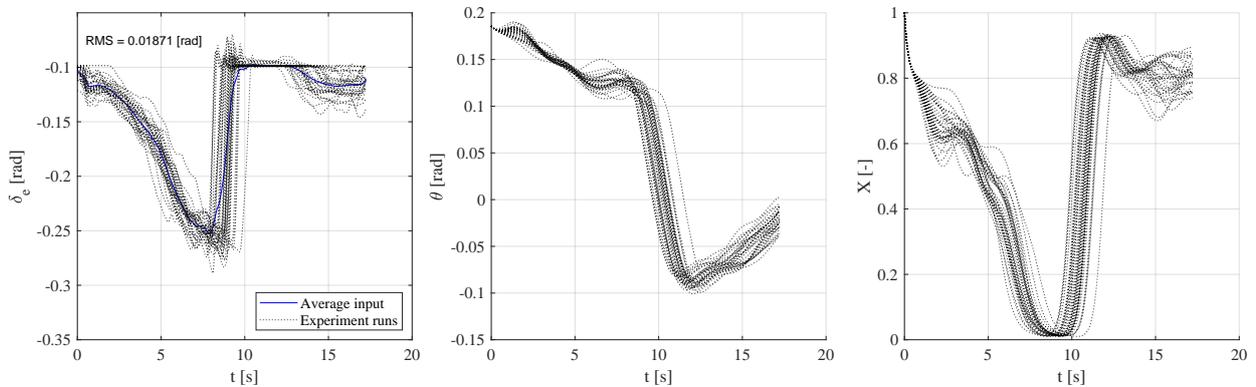


Figure F-12: Elevator inputs for all baseline runs from Subject 6, together with the pitch angle and flow separation point.

Subject 7

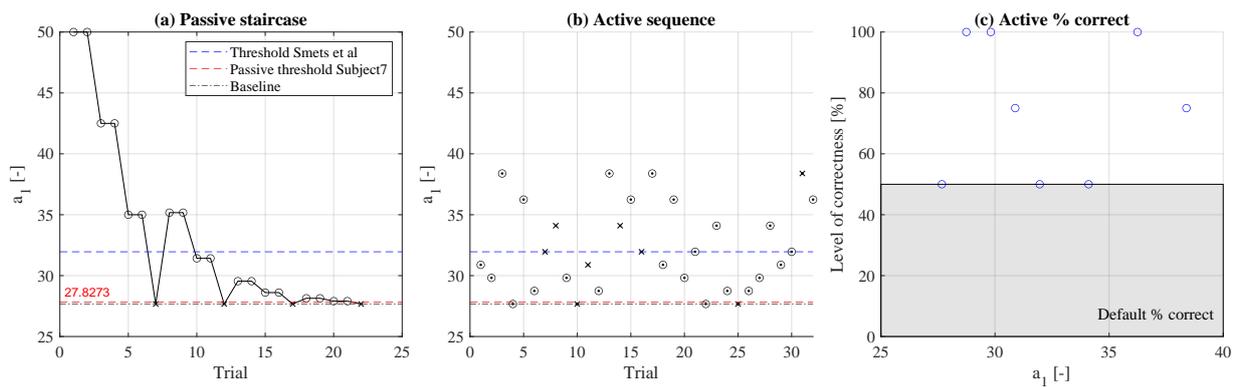


Figure F-13: Results of both the passive and active experiment for Subject 7, presented together with the percentages correct for the active experiment.

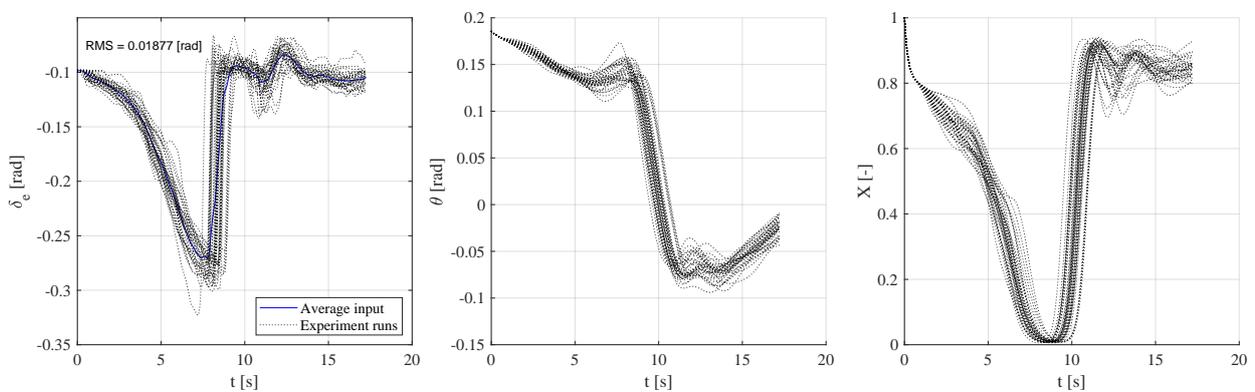


Figure F-14: Elevator inputs for all baseline runs from Subject 7, together with the pitch angle and flow separation point.

Subject 8

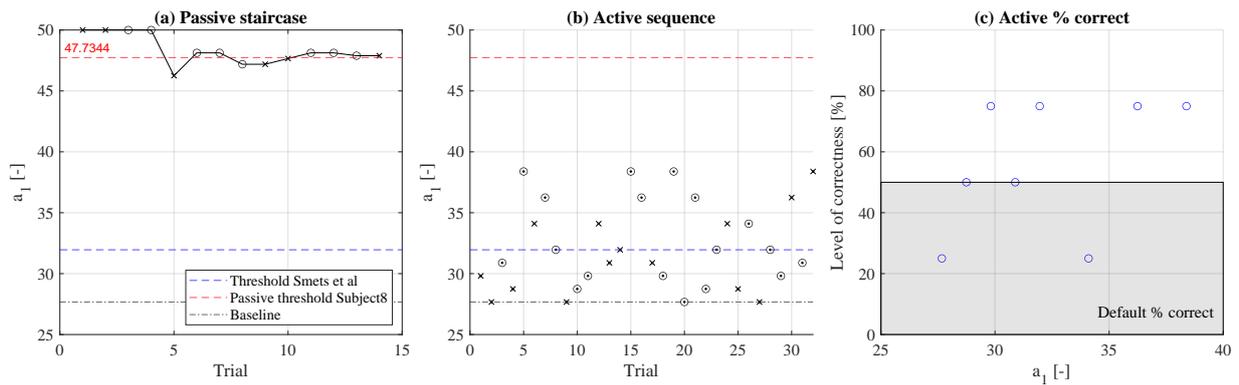


Figure F-15: Results of both the passive and active experiment for Subject 8, presented together with the percentages correct for the active experiment.

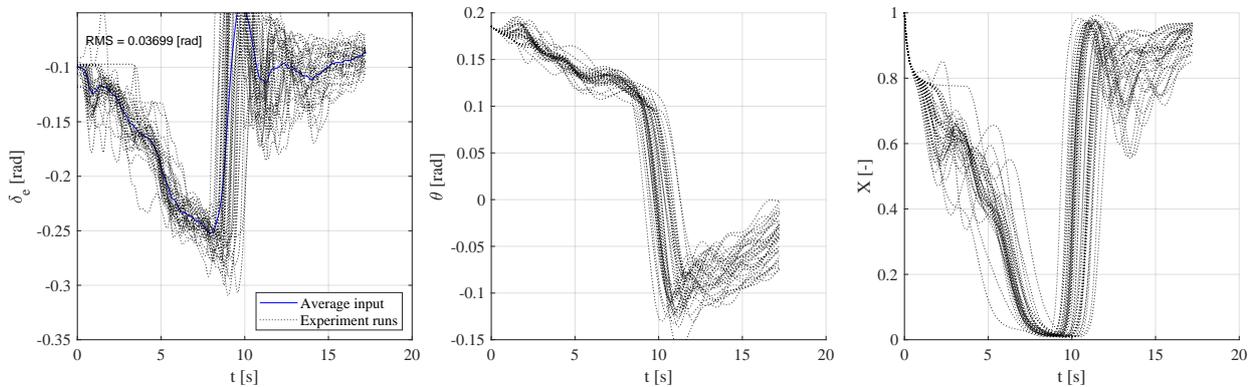


Figure F-16: Elevator inputs for all baseline runs from Subject 8, together with the pitch angle and flow separation point.

Subject 9

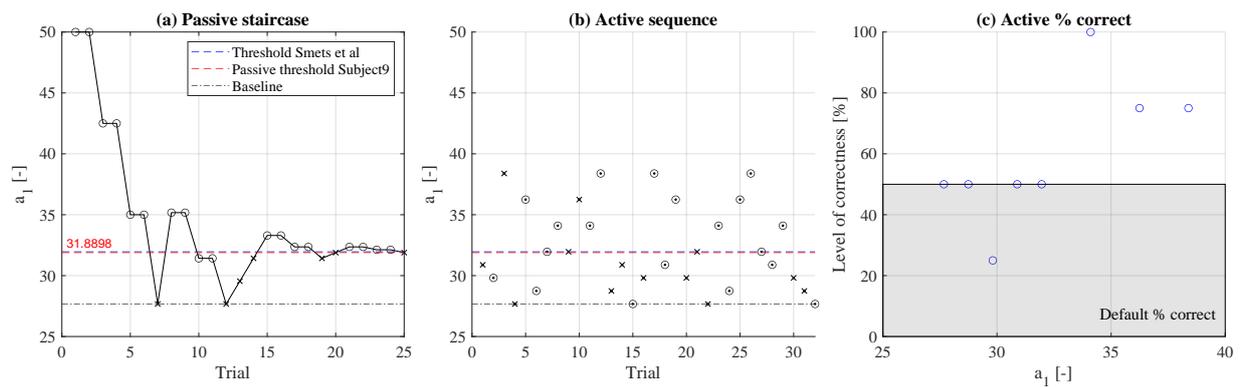


Figure F-17: Results of both the passive and active experiment for Subject 9, presented together with the percentages correct for the active experiment.

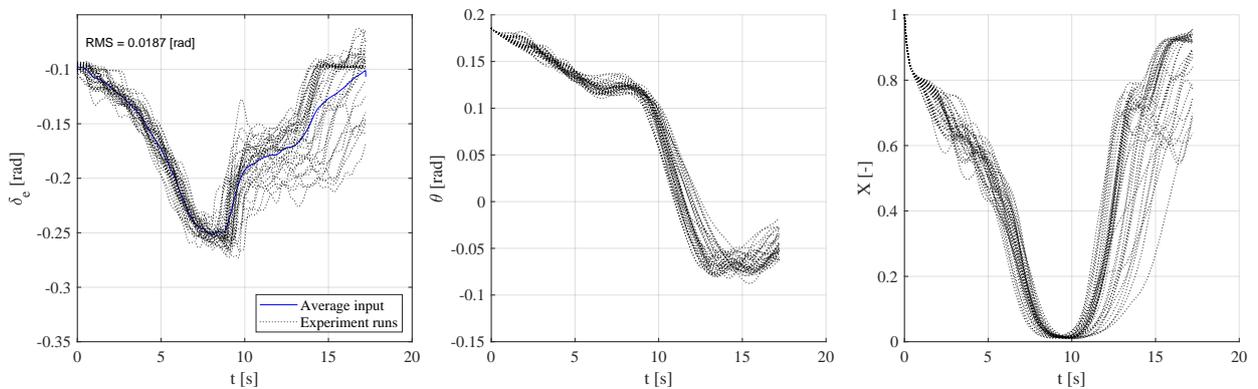


Figure F-18: Elevator inputs for all baseline runs from Subject 9, together with the pitch angle and flow separation point.

Subject 10

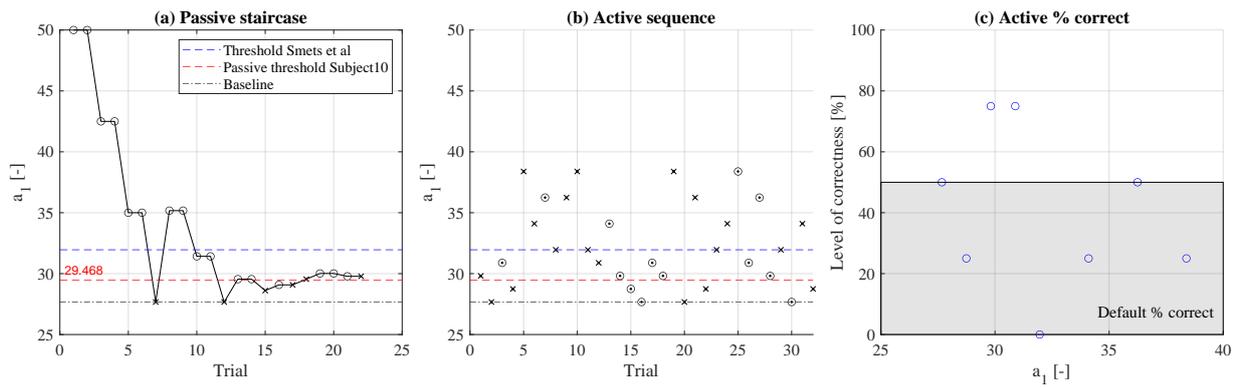


Figure F-19: Results of both the passive and active experiment for Subject 10, presented together with the percentages correct for the active experiment.

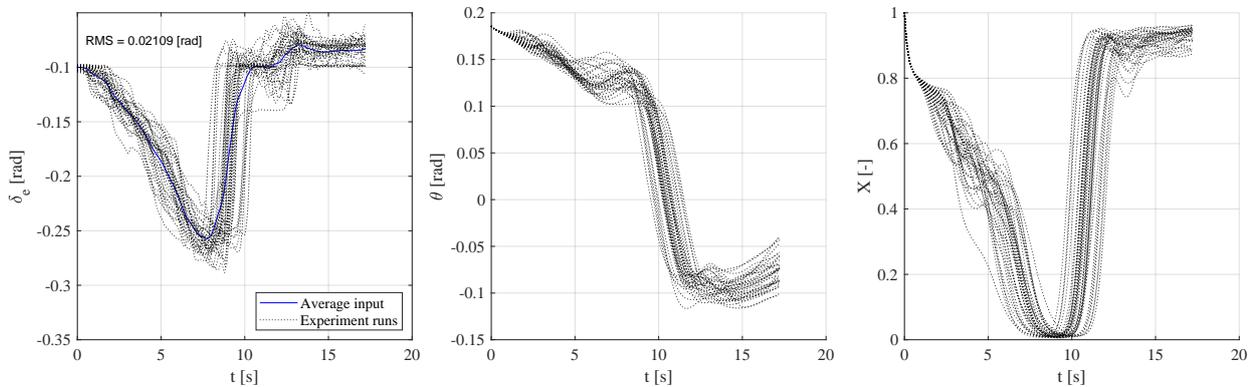


Figure F-20: Elevator inputs for all baseline runs from Subject 10, together with the pitch angle and flow separation point.

Subject 11

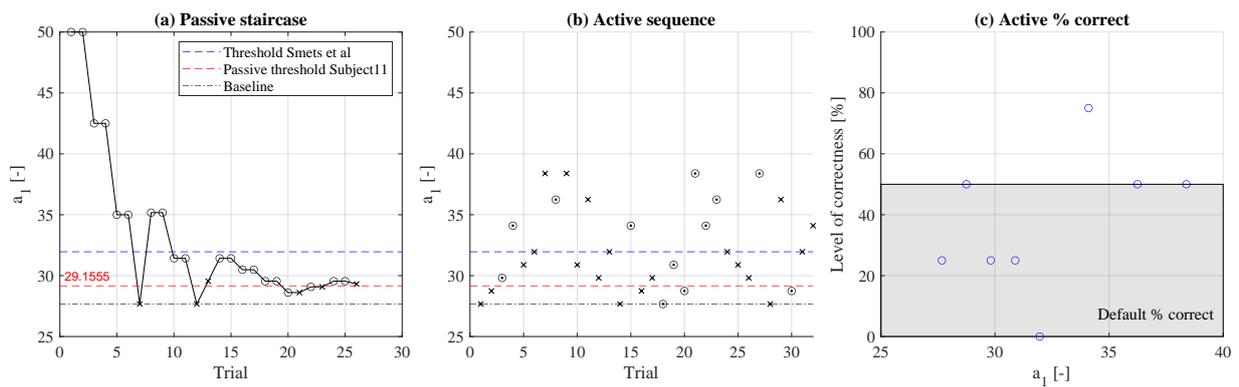


Figure F-21: Results of both the passive and active experiment for Subject 11, presented together with the percentages correct for the active experiment.

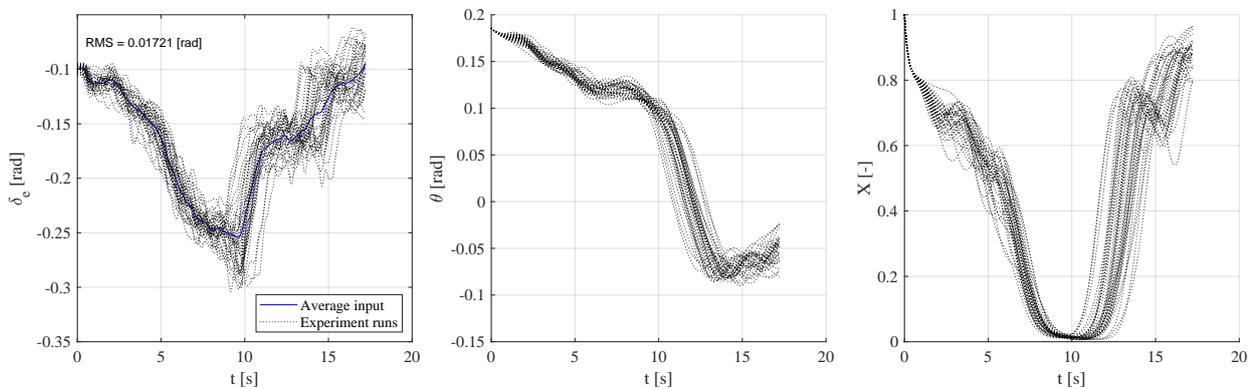


Figure F-22: Elevator inputs for all baseline runs from Subject 11, together with the pitch angle and flow separation point.

Subject 12

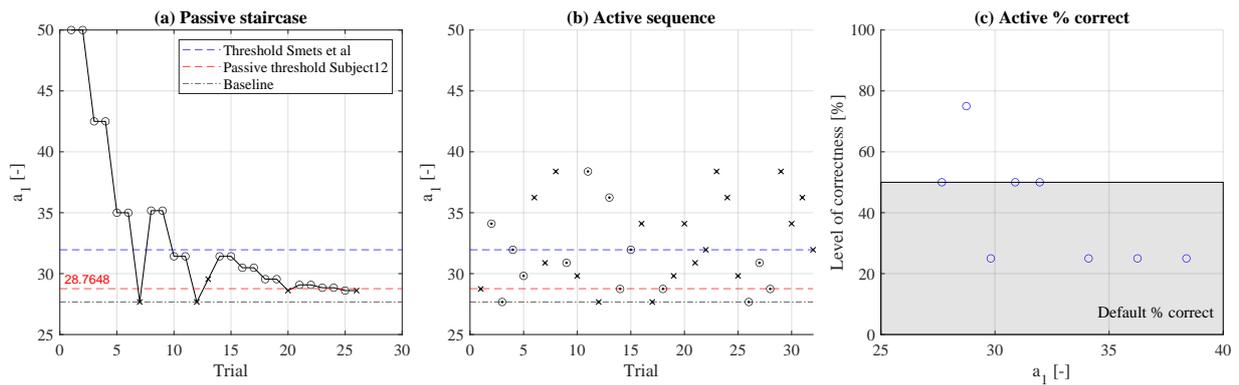


Figure F-23: Results of both the passive and active experiment for Subject 12, presented together with the percentages correct for the active experiment.

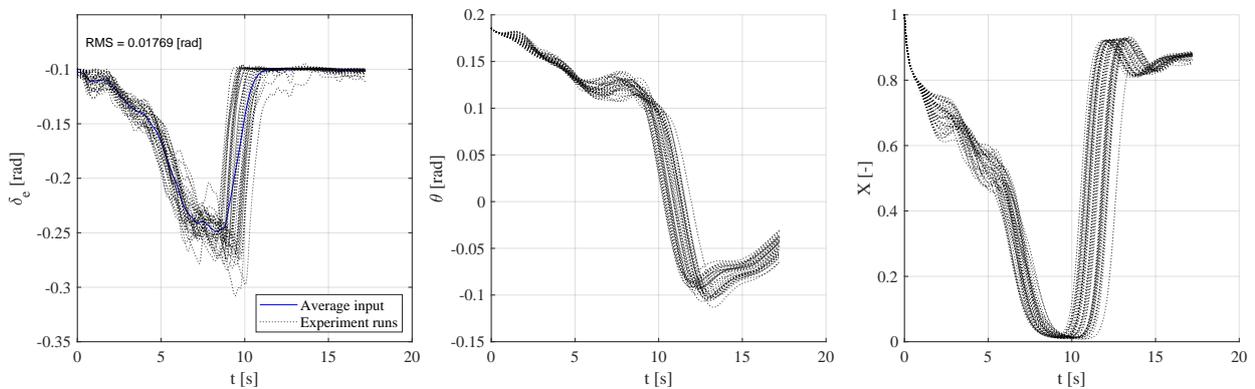


Figure F-24: Elevator inputs for all baseline runs from Subject 12, together with the pitch angle and flow separation point.

Subject 13

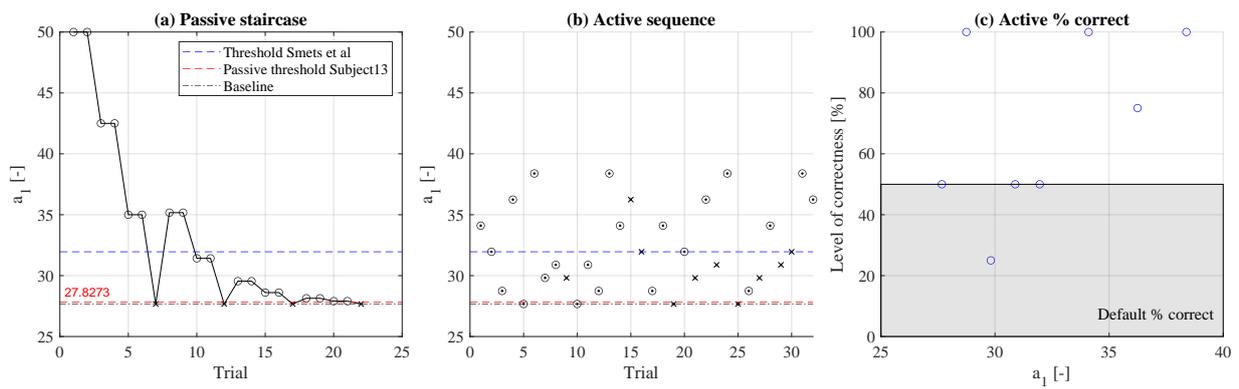


Figure F-25: Results of both the passive and active experiment for Subject 13, presented together with the percentages correct for the active experiment.

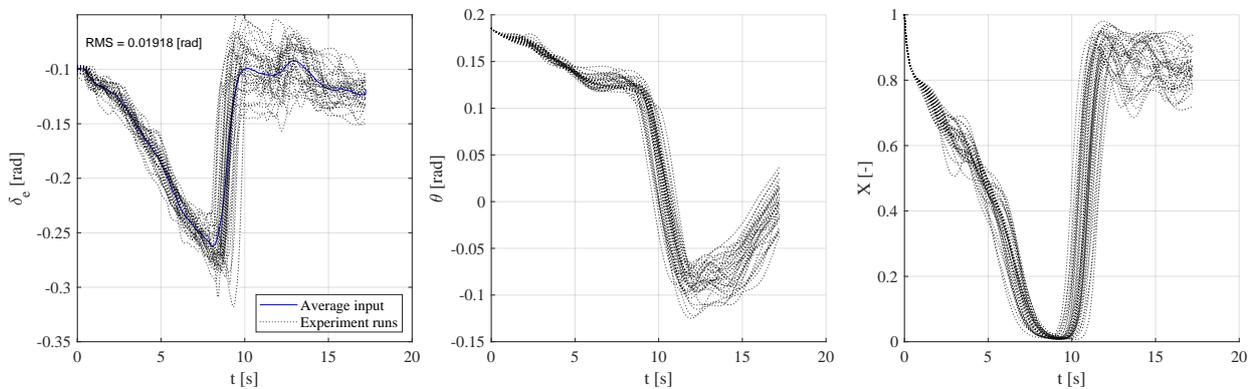


Figure F-26: Elevator inputs for all baseline runs from Subject 13, together with the pitch angle and flow separation point.

Subject 14

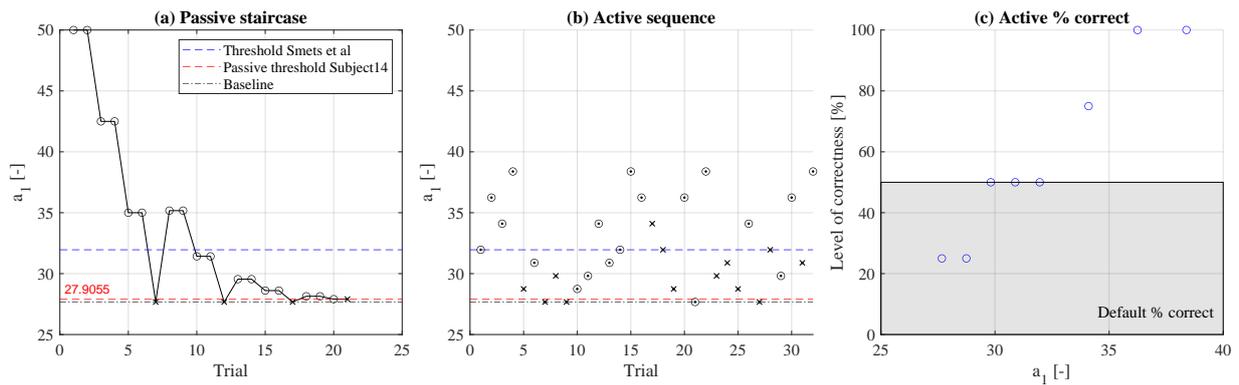


Figure F-27: Results of both the passive and active experiment for Subject 14, presented together with the percentages correct for the active experiment.

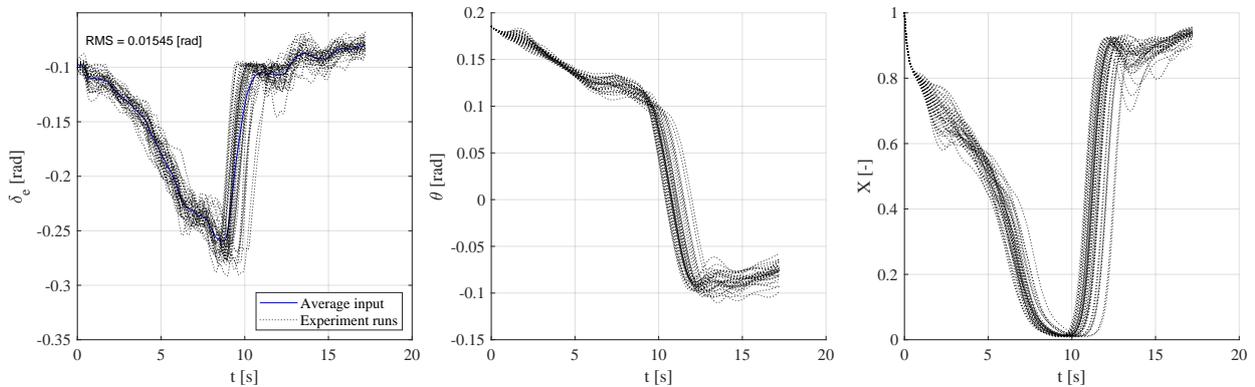


Figure F-28: Elevator inputs for all baseline runs from Subject 14, together with the pitch angle and flow separation point.

Subject 15

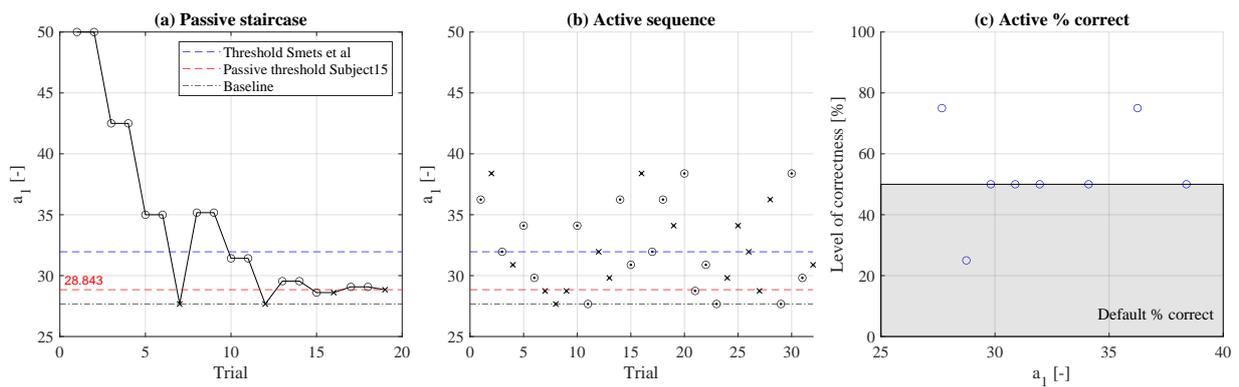


Figure F-29: Results of both the passive and active experiment for Subject 15, presented together with the percentages correct for the active experiment.

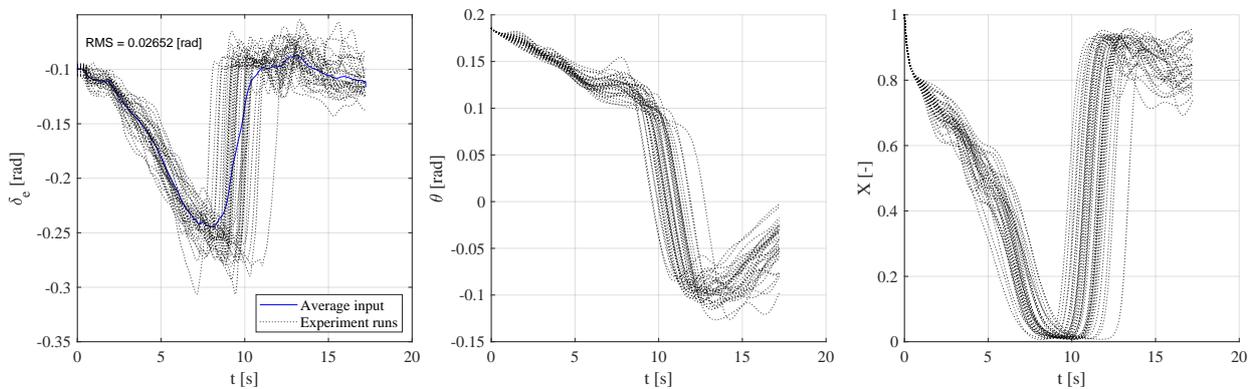


Figure F-30: Elevator inputs for all baseline runs from Subject 15, together with the pitch angle and flow separation point.

Subject 16

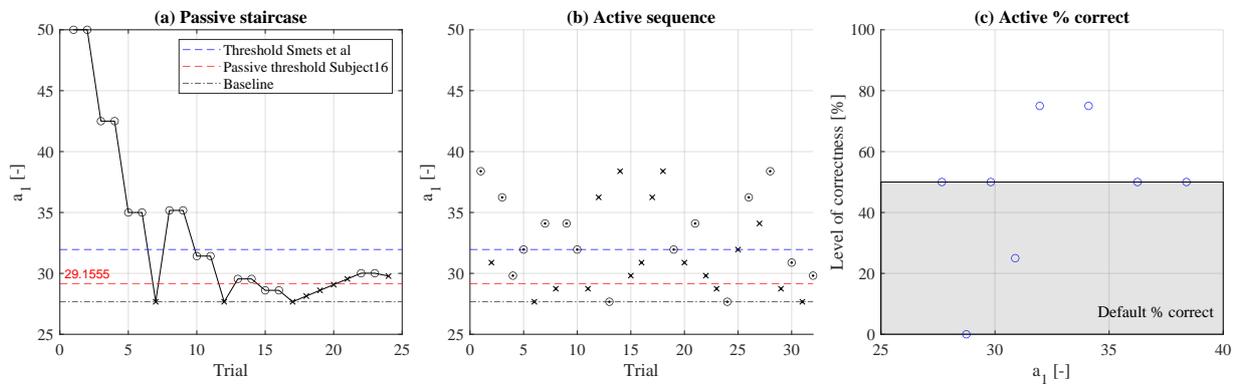


Figure F-31: Results of both the passive and active experiment for Subject 16, presented together with the percentages correct for the active experiment.

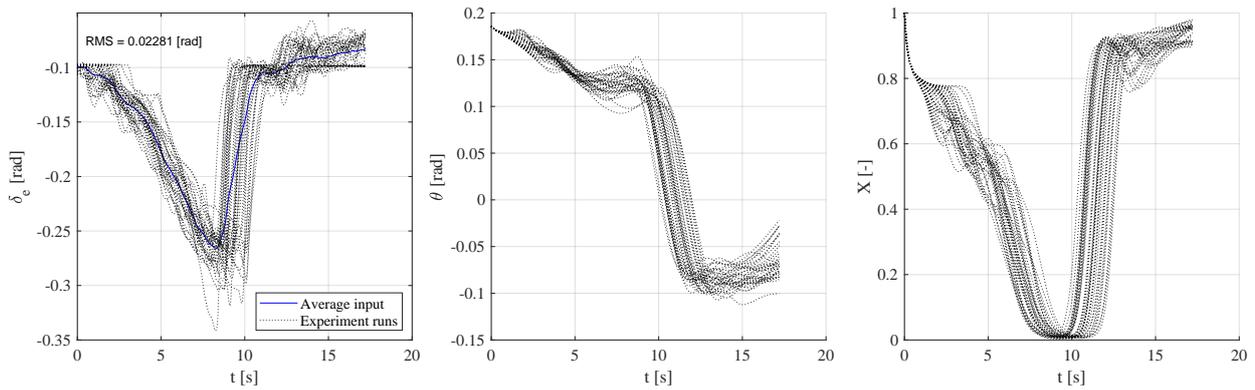


Figure F-32: Elevator inputs for all baseline runs from Subject 16, together with the pitch angle and flow separation point.