



ON THE CHALLENGES OF SURROGATE-BASED AERODYNAMIC SHAPE OPTIMISATION

Juan Rasines¹ & Dries Verstraete¹

¹School of Aerospace, Mechanical and Mechatronic Engineering, The University of Sydney, Australia

Abstract

The increasing use of optimisation during early design phases raises challenges associated with computational efficiency and simulation fidelity. Surrogate-based optimisation is a promising tool for designers that seek to understand multi-objective trade-offs and produce optimal solutions to problems involving high-fidelity aerodynamic simulations, in a short amount of time. Its potential reduction in the number function evaluations merits the application of surrogate modelling, yet there are associated challenges, including the occurrence of simulation failures which is rarely addressed in optimisation research. This work investigates the effect of simulation failure handling strategies in surrogate-based evolutionary optimisation, applied to airfoil design optimisation using low- and high-fidelity aerodynamic simulation. Convergence studies on a supercritical airfoil design problem demonstrate the importance of selecting the correct simulation failure handling strategy when using surrogate-based optimisation algorithms. Moreover, the characterisation of simulation failures in the design space demonstrate differences in occurrence and distribution between low- and high-fidelity simulation (Euler and Reynolds-Averaged Navier Stokes codes, respectively), which is critical to inform algorithm design and tailoring to the desired simulation fidelity for efficient surrogate-based optimisation.

Keywords: Aerodynamic shape optimisation, airfoil design, multi-objective optimisation, simulation failures, surrogate-based optimisation, high-fidelity simulation

1. Introduction

Real-world engineering design problems involving simulation-based optimisation are subject to simulation failures [1, 2, 3, 4]. In aerodynamic shape optimisation, infeasible geometries and poor quality meshes can cause solutions to not convergence or lead to numerical instability in Computational Fluid Dynamics (CFD) codes [5]. Such simulation failures impose additional computational burden and add complexity to design problems that already exhibit other challenges like high-dimensionality in the input variables, multi-objective trade-offs, highly-constrained design spaces, and computationally expensive high-fidelity simulation [6, 7, 8, 9, 10, 11]. Few global optimisation algorithms can effectively solve these problem challenges within reasonable computational timeframes. Surrogate-based optimisation is a promising alternative to traditional gradient-free optimisers, as the number of required function evaluations can be drastically reduced compared to evolutionary algorithms [9]. Surrogate-based methods typically use regression models to approximate the underlying functions and reduce the number of calls to the expensive simulation code [7, 12]. The ability to optimise at a reduced computational expense is of particular interest given the growing trend in higher-fidelity simulation during the early design stages in the aerospace industry [7, 13]

Despite the potential benefits, surrogate-based optimisation algorithms face many unresolved challenges when applied to real-world engineering problems [7, 9, 12]. In particular, the occurrence and effect of simulation failures on surrogate modelling and algorithm convergence has received limited attention in literature, especially in the context of aerodynamic shape optimisation. Early work by Tenne and Armfield [14] proposed a surrogate-based memetic algorithm capable of handling simulation failures, in the context of airfoil shape optimisation for a small business jet. A penalty method approach was used to modify the surrogate landscape, increasing the objective function value in the

vicinity of undefined evaluations. Baert et al. [2] incorporated a probabilistic Support Vector Machine (SVM) model within the MINAMO framework [15] to classify failure regions in the design space. The proposed methodology was successfully used to optimise a highly-constrained transonic compressor [2]. Similarly, Wilke [3] trained a surrogate on simulation failure information to predict a probability of failed evaluation, as a pre-filtering step in the infill criterion. High-fidelity optimisation of helicopter rotor blades was performed using the surrogate-based framework, achieving up to 78% reduction in computational resources when applying the pre-filtering surrogate, compared to classical simulation failure handling [3]. Other work by Arsenyev et al. [1] built a Gaussian Process (GP) regressor on simulation failure information by considering the Euclidean distance to the nearest successful evaluation to limit oscillations in the surrogate model. A multi-disciplinary optimisation of a Low-Pressure Turbine (LPT) vane cluster was demonstrated using high-fidelity Reynolds-averaged Navier Stokes (RANS), with the incorporated simulation failure handling surrogate.

Various simulation failure handling strategies for surrogate-based optimisation have been proposed in literature, including its application towards expensive, high-fidelity aerodynamic shape optimisation. Nevertheless, the tools proposed in literature have only been evaluated on single-objective optimisation problems, with a relatively small percentage of simulation failures. Additionally, the effects of solver fidelity on simulation failure in surrogate-based optimisation is yet to be characterised. This work aims to quantify the effectiveness of simulation failure handling (SFH) strategies for surrogate-based evolutionary optimisation, as applied to the aerodynamic shape optimisation of an airfoil subject to a high degree of simulation failures. A baseline surrogate-based constrained multi-objective evolutionary algorithm is considered, and six methodologies to handle simulation failures, as modified from literature, are investigated. The effect of simulation failure handling on algorithm convergence is evaluated on a constrained, multi-objective airfoil design problem, using low- and high-fidelity aerodynamic solvers, MSES [16] and SU2 [17], respectively. Results are discussed with respect to the applicability of solving complex aerodynamic shape design problems, and the viability of surrogate-based optimisation to handle simulation failures.

The following sections describe the methodology developed to evaluate the aerodynamic shape optimisation problem. This includes; problem definition, airfoil parametrisation, and aerodynamic evaluation with a low- and high-fidelity solvers. The surrogate-based optimisation framework is introduced alongside with six simulation failure handling strategies investigated. Results are presented for the convergence studies conducted, highlighting key take-aways for surrogate-based optimisation algorithm design. Lastly, simulation failures are characterised in the design space for both the low- and high-fidelity aerodynamic solvers.

2. Aerodynamic Shape Optimisation Problem Methodology

The selected airfoil design problem is derived from the development of the NASA supercritical airfoil family [18]. Specifically, the NASA third phase design methodology is applied to an operating regime subjected to compressibility effects. This problem is particularly interesting to optimise given the high degree of simulation failures. These are expected as the operating conditions will result in weak and strong shock effects, as well as complex nonlinear flowfields that are difficult to (re)-solve. The SC(2)-0412 airfoil is used as the baseline, which has a design lift coefficient of $C_L = 0.4$ at a Reynolds number (Re) of $30 \cdot 10^6$ with a maximum thickness of 12% of the chord c . The design objectives include maximising the weighted lift-to-drag at the design coefficient across the Mach number operating range (Equation 1) and maximising the drag-divergence Mach number M_{dd} (Equation 2).

$$\overline{L/D} = \sum_{i=0}^{N_{fc}} \frac{\omega \cdot C_L}{C_d} \Big|_i \quad \text{for } M \in [0.6, 0.8] \quad (1)$$

$$M_{dd} = M \Big|_{\partial C_d / \partial M = 0.1} \quad \text{for } M \in [0.6, 0.82] \quad (2)$$

The partial derivative of the drag coefficient with respect to Mach number requires a smooth continuous function, which is achieved through quadratic interpolation of the evaluated points (a minimum of

three non-failed points are required). The objective weightings $\omega_i = 1$ are set equal for the Mach number range. Specifically, the evaluated Mach numbers are $M = [0.6, 0.68, 0.74, 0.78, 0.8]$, which equates to five flight conditions ($N_{fc} = 5$).

The parametrisation of the airfoil is achieved with Bezier spline curves for the top and bottom surfaces, respectively. A total of ten uniformly spaced control points are used to define both Bezier curves, which are affine summations of Bernstein polynomials [19]. The parametrisation scheme is visualised in Figure 1 for the SC(2)-0412 section, which defines some of the geometric constraints used. A total of 16 geometrical constraints are applied to obtain reasonable airfoil geometries, as summarised in Table 1.

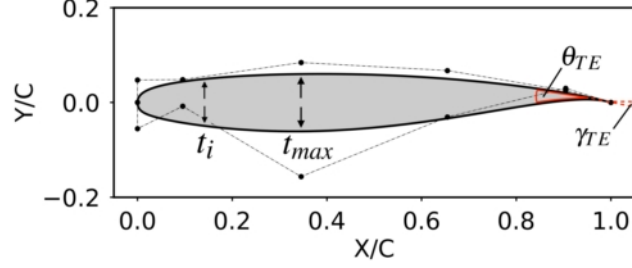


Figure 1 – Bezier parametrisation of the SC(2)-0412 section with ten shape variables, showing control points on upper and lower surfaces. Key geometric features relating to constraints are labelled.

Table 1 – Summary of geometrical constraints applied to the airfoil problem.

Constraint	Number of constraints	Expensive	Description
CR	1	N	Crossover between upper and lower surfaces
t_i	7	N	Minimum airfoil thickness at chord location c_i
$N_{cvr,top}$	1	N	Number of curvature reversals on the upper surface
$N_{cvr,bot}$	1	N	Number of curvature reversals on the bottom surface
θ_{TE}	2	N	Trailing edge angle with the horizontal (degrees)
γ_{TE}	2	N	Trailing edge boat-tail angle (degrees)
$C_L _{climb}$	1	Y	Minimum lift coefficient requirement at climb conditions
Total	16		

Here, the crossover constraint CR is the negative minimum thickness between the upper and lower surfaces, so that if the surfaces overlap at any chord location, CR becomes positive. The airfoil thickness t_i at chord locations $c_i \in [0.005c, 0.01c, 0.05c, 0.2c, 0.8c, 0.9c, 0.95c]$ is used to maintain a minimum thickness $t_{min|i}$ relative to the baseline SC(2)-0412 section. Two angles (γ_{TE} and θ_{TE}) are used to control the trailing edge shape, and the bounds for the four corresponding constraints are selected to avoid excessively thin and highly-cambered trailing edges. This is important as supercritical sections are typically aft-loaded [20]. Limiting the number of curvature reversals ($N_{cvr,top}$ and $N_{cvr,bot}$) ensures that excessive local oscillations are not present in the optimal shapes. This constraint is evaluated by counting the number of sign reversals in the derivative of the local spline curvature with respect to the arclength, for the top and bottom surfaces individually. The maximum airfoil thickness t_{max} must be larger than 12% of the chord, to match the specifications of the SC(2)-0412 section. Lastly, one computationally expensive constraint is applied, which is a minimum lift coefficient requirement of $C_L \geq 0.9$ for an angle of attack of 5 degrees at climb conditions ($Re = 8.8 \cdot 10^6$ and $M = 0.22$). The optimisation problem is thus defined in Equation 3, where $\mathbf{x}_L$ and $\mathbf{x}_U$ are the lower and upper bounds of the design space, respectively. The Bezier airfoil bounds are determined by fitting the University of Illinois Urbana-Champaign (UIUC) airfoil database with the Bezier parametrisation and extracting

the minimum and maximum values for each control point.

$$\begin{aligned}
 &\text{Minimize} && \left\{ \overline{L/D}, \quad M_{dd} \right\} \\
 &\text{S.t.} && CR \leq 0 \\
 &&& t_{min|i} - t_i \leq 0 \quad i = 0, \dots, 6 \\
 &&& 0.12 - t_{max} \leq 0 \\
 &&& 1 - N_{cvr,top} \leq 0 \\
 &&& 2 - N_{cvr,bot} \leq 0 \\
 &&& \theta_{TE} - 25 \leq 0 \\
 &&& 0 - \theta_{TE} \leq 0 \\
 &&& \gamma_{TE} - 12.5 \leq 0 \\
 &&& 5 - \gamma_{TE} \leq 0 \\
 &&& 0.9 - C_L|_{climb} \leq 0 \\
 &&& \mathbf{x}_L \leq \mathbf{x} \leq \mathbf{x}_U
 \end{aligned}$$

2.1 Low fidelity simulation

The MSES code from the Massachusetts Institute of Technology (MIT) [16] is a low fidelity solver capable of capturing transonic effects. MSES solves the Euler equations with a two-equation integral boundary layer solver, almost identical to that of XFOIL [21]. The accuracy and general agreement of MSES with experimental data is recognised for subsonic and transonic compressible flows [22, 23, 24, 25], similarly to XFOIL for subsonic flows [13, 26, 27]. Due to the similar viscous-drag correction methodologies of XFOIL and MSES, is it expected that the simulation failure-prone tendencies of the former [14, 27] are also observed for the latter. An additional source of convergence difficulty that features in the MSES code derives from the angle of attack specification in the initial inviscid panel solution. This step defines the iso-potentials used to discretise the grid generation for the numeral evaluation of the Euler equations. In practice, experience shows that the convergence failure rate of MSES for transonic flows is pronouncedly higher than for XFOIL in subsonic flows, including very low Reynolds numbers.

2.2 High fidelity simulation

To compare the occurrence and frequency of simulation failures in aerodynamic solvers, a comparison against MSES is made by incorporating a high-fidelity CFD solver within the surrogate-based optimisation framework. The Stanford University Unstructured (SU2) multi-physics code [17] is used to solve the compressible RANS equations using a Green-Gauss cell discretisation for spatial gradients. The flow inviscid terms are discretised with a Flux Difference Splitting (FDS) upwind scheme. Slope-limiting is set using the Venkatakrishnan method, and Flexible Generalised Minimum Residual (FGMRES) is used as the linear solver with LU-SGS preconditioning [17]. The Spalart-Allamaras one-equation BCM transition turbulence model (SA-BCM) [28] is chosen through experimentation and solver validation at the airfoil design conditions. The minimum residual is set to $1 \cdot 10^{-6}$ for the density field and a maximum of 15,000 iterations is allowed. An O-grid mesh is generated using the automated meshing tool, Construct2d [29]. The selection of the mesh structure eliminates the need to re-generate the geometry for different angles of attack. The mesh domain boundaries are at 20 chord-lengths from the airfoil. The airfoil surface is discretised with 501 points, and the boundary layer is captured with a first-layer height of $y^+ = 0.5$. The mesh discretisation is selected through a convergence study conducted in ref. [30], where the current settings result in an acceptable trade-off between computational runtime and variation in aerodynamic coefficients.

2.3 Aerodynamic solver validation

A comparison of MSES and SU2 against experimental wind tunnel data is conducted for the SC(2)-0710 airfoil. This section is evaluated over the baseline SC(2)-0412 due to availability of experimental results in ref. [31]. Both aerodynamic solvers are evaluated for a Mach number sweep between $M \in [0.5, 0.8]$ at a design lift coefficient of $C_L = 0.7$ and a Reynolds number of $Re = 40 \cdot 10^6$, which represent very similar operating conditions to the design problem to be optimised. MSES is evaluated with free-transition and a critical intensity factor of $N_{crit} = 8.1$ to match the turbulence intensity factor

of $TI = 0.1\%$ used in SU2, as per the wind tunnel conditions [31]. The rest of the solver settings are as previously described. Comparison to experimental wind tunnel results in Figure 2 demonstrates a general under-prediction of drag by MSES. SU2 with the SA-BCM model demonstrates strong agreement with experimental drag measurements up to Mach 0.7, after which the onset of wave drag from compressibility effects is over-predicted. Subsequently, the drag-divergence Mach number (shown as vertical lines in Figure 2) is under-predicted. A similar trend is observed for MSES, however the onset of wave drag initially appears more aligned due to the drag offset. The drag-divergence Mach numbers are evaluated as $M_{dd} = 0.764, 0.766, 0.780$ for MSES, SU2, and experimental data, respectively. As seen, both low- and high-fidelity codes fail to capture the localised reduction in drag around Mach 0.77 seen in the experimental results. The onset of wave drag is thus over-predicted, however, a small discrepancy in the M_{dd} predictions is obtained, with a relative error of -2.1% and -1.7% for MSES and SU2, respectively. The observed trends from both solvers agree well with experimental data, and enable sufficient confidence to conduct numerical optimisation using the aerodynamic predictions obtained from MSES and SU2.

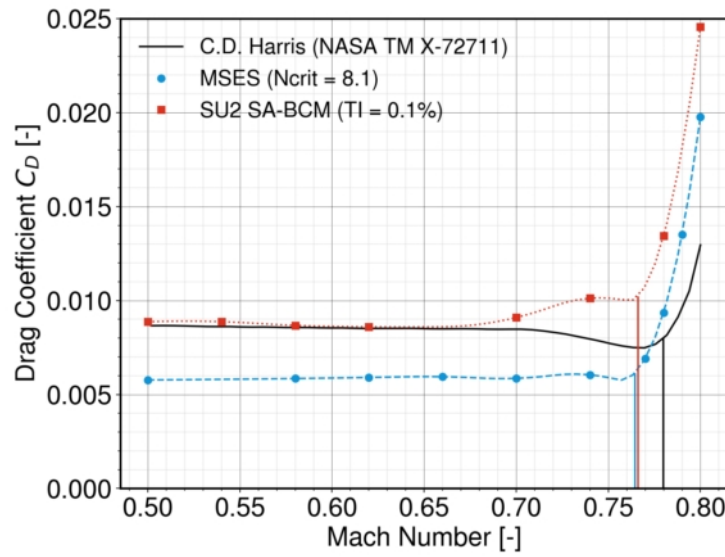


Figure 2 – Experimental validation of MSES and SU2 against wind tunnel data [31] for the SC(2)-0710 airfoil at $C_L = 0.7$, $Re = 40 \cdot 10^6$. Vertical lines represent the calculated drag-divergence Mach number.

3. Surrogate-based Optimisation Methodology

A simple Surrogate-Based Evolutionary Algorithm (SAEA) is considered in this work to solve the constrained multi-objective problem, as the baseline for simulation failure handling benchmarking. More specifically, the traditional Non-dominated Sorting Genetic Algorithm (NSGA-II) [32] is augmented with surrogates using regression models for each objective and constraint individually. The framework and settings of the SAEA are as follows:

1. **Initialisation:** The initial Design of Experiments (DoE) is generated using Latin Hypercube Sampling (LHS) with a size of 10D, where $D = 10$ is the number of design variables. These designs are evaluated expensively using simulation codes MSES and SU2. Different LHS seeds are pre-run and loaded *a priori* to each optimisation run for consistent comparisons between simulation failure handling strategies and aerodynamic solvers.
2. **Surrogate Management:** Radial Basis Functions (RBF) are used as surrogates to regress each objective and constraint individually. All RBF models use a cubic kernel and linear tail, as they perform well for a range of problem landscapes [33]. The regression models are initialised using the DoE evaluations and newly evaluated points are used at each generation to update the surrogate models in an online optimisation approach.

3. **Evolution:** The traditional NSGA-II evolutionary algorithm is evolved for ten generations with a population size of 60 and offspring size of 300. Convergence is accelerated by evaluating NSGA-II on the computationally-cheap surrogate predictions. The final evolved population after ten generations survives to the next stage – the infill point selection.
4. **Infill Criterion:** To update the surrogate models efficiently towards the global optima, one infill point is selected per iteration for expensive evaluation using simulation. The Hypervolume Improvement (HVI) [34] (see Section 5.) is calculated between the NSGA-II population and the archived population. The infill point is the solution with the maximum HVI score, which represents the point with the greatest convergence-gain over the current optimal population.
5. **Termination:** A total infill computational budget of 50D is used to terminate the algorithm, which is 300 times smaller than the number of function evaluations used to approximate the Pareto Front (PF) in Figure 4 – (a), using the standard constrained NSGA-II formulation. The online optimisation infill process uses 83% of the allowed computational budget, which is conducted serially. For the high fidelity SU2 optimisation runs, the budget is reduced by half to 25D due to the long computational run-times. This particular problem requires eight full CFD evaluations per geometry to calculate the objectives and constraints, which equates to over one hour of computation on a 32-core machine.

4. Simulation Failure Handling Methodology

Six strategies to handle failed evaluations are investigated by modifying the baseline SAEA, of which the first four are visualised in Figure 3. For the considered airfoil shape optimisation problem, simulation failures can occur for both objectives and the expensive lift constraint, which are set to 'not-a-number' (NaN). A measure of failed evaluations can be defined by the Simulation Failure Rate (SFR), which counts the proportion of NaN's in the objectives and constraints for a population, as presented in Equation 3.

$$\text{SFR} = \frac{|\{\mathbf{p} \in \mathbb{P} \mid \text{failed}(\mathbf{f}, \mathbf{g})\}|}{|\mathbb{P}|} \quad (3)$$

Where $\mathbf{p}$ is an individual in the population set $\mathbb{P}$ and $(\mathbf{f}, \mathbf{g})$ are the corresponding objective and constraint vectors, respectively. The function $\text{failed}(\mathbf{x}, \mathbf{y})$ indicates whether any components of the vectors $\mathbf{x}$ and $\mathbf{y}$ have any NaN's present. Strategy one (SFH1) can be considered the 'naive' approach

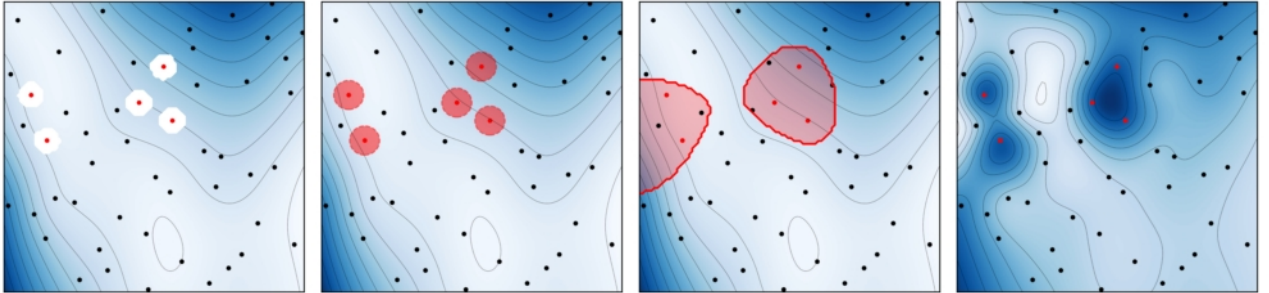


Figure 3 – Four simulation failure handling strategies in order SFH1–4 left-to-right: showing DoE with black points, simulation failures with red points, and contours for the surrogate predictions. Predicted failure-prone areas are shown in red shaded areas for SFH2 and SFH3. Light colours represent the minima of the Branin-Hoo test function, and dark colours the high-objective regions. White patches indicate missing objective information due to failed evaluations.

and consists only of filtering out failed evaluations from initial surrogate training after the DoE, and through subsequent infill point evaluations. Strategy 2 (SFH2) extends upon SFH1 by applying a trust region around previously failed evaluations to prevent the algorithm from evaluating points close to failure regions in the design space. SFH2 is used to pre-select the NSGA-II population before infill point selection, where the trust region threshold R_{crit} is set experimentally to 5% of the Euclidean distance of the design space. If no infill candidates lie outside the trust regions, R_{crit} is decreased

iteratively by 10% until at least one infill point exists. Strategy 3 (SFH3) uses a probabilistic SVM classification surrogate, equivalent to Baert et al. [2]. The SVM surrogate is trained on simulation failures to pre-filter points likely to return a failed evaluation before infill point selection, similarly to SFH2. Strategy four (SFH4) can be considered the 'classical' penalty approach, where large objective values are imputed for failed evaluations. The optimiser filters out these points due to their poor fitness ranking, at the expense of the surrogate objective landscapes becoming distorted. For each failed evaluation infill point, the surrogate objective prediction is penalised by a factor of 1.1 and imputed into the archive, before updating the surrogate models.

Strategy five (SFH5) is identical to SFH1 besides in augmenting the standard LHS initialisation. Evolution is conducted to obtain a prescribed size of geometrically-feasible solutions before conducting expensive evaluation of the LHS population. Geometric constraints for this problem are analytical and considered 'computationally cheap' to evaluate. This strategy aims to reduce the occurrence of initial simulation failures, whilst maintaining as much diversity in the design space as possible. Geometric filtering has the added benefit of automatically bypassing the redundant regions of the search space, such that the surrogate models are trained to be more accurate in the regions of interest. Lastly, strategy six (SFH6) extends SFH5 by exploiting computationally-cheap geometric constraint information throughout the entire optimisation process. Specifically, the surrogate predictions for geometric constraints are replaced with the real-evaluated constraint values during the internal offspring evolutions in each optimisation iteration. The purpose of SFH6 is to reduce surrogate accuracy biases for constraints that need not be regressed due to the availability of fast, exact constraint information.

5. Optimisation Convergence Methodology

The Hypervolume indicator metric [34] is used to calculate the HVI in the infill point selection during the surrogate-based optimisations. It is also used to calculate the Hypervolume Coverage (HVC) of an optimal set of solutions against a pre-determined Pareto Front. The HVC measures the level of coverage of a set of Pareto-optimal solutions with respect to the best-known Pareto front ($\mathbf{PF}^*$) of the problem. The Hypervolume Improvement (HVI) is defined as per Equation 4 and the HVC as per Equation 5.

$$\text{HVI} = HV(\mathbb{P}, \mathbf{r}) - HV(\mathbf{PF}^*, \mathbf{r}) \quad (4)$$

$$\text{HVC} = \frac{HV(\mathbb{P}, \mathbf{r})}{HV(\mathbf{PF}^*, \mathbf{r})} \quad (5)$$

Where HV is the Hypervolume of a population $\mathbb{P}$ with respect to a nadir point $\mathbf{r}$. The HVC is expressed as a fraction ($\text{HVC} \in [0, 1]$), therefore a value of $\text{HVC} = 1$ means the population $\mathbb{P}$ has a perfect coverage, and is identical, to the best-known Pareto front ($\mathbf{PF}^*$). For the considered objectives, in the context of minimisation, the nadir point is approximated from MSES runs as $\mathbf{r} = [-45.171, -0.7885]$.

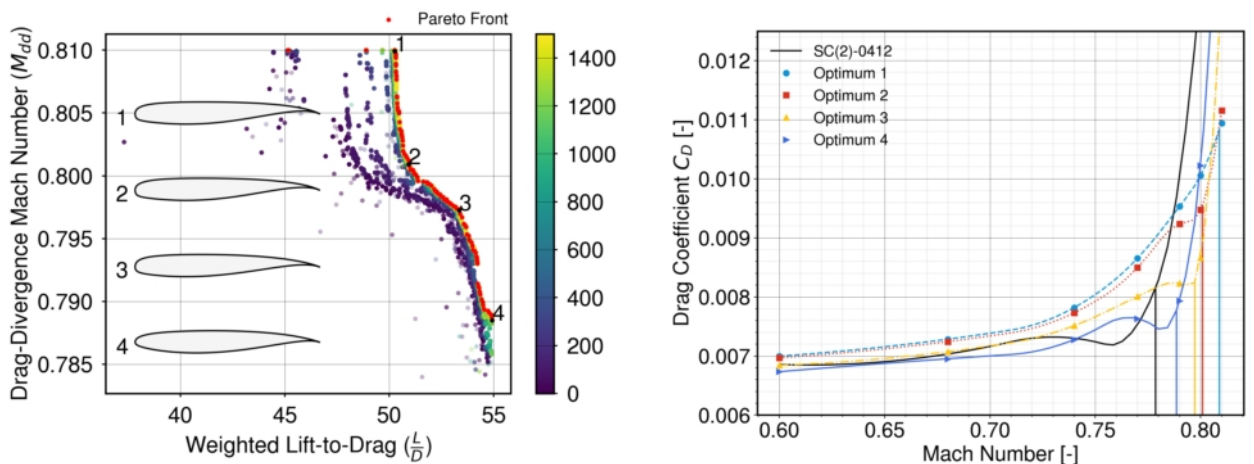
6. Results

The following section presents the convergence studies conducted on the supercritical airfoil design problem using the low- and high-fidelity aerodynamic solvers, MSES and SU2, respectively. Firstly, the approximated Pareto Front for MSES is obtained and the aerodynamic performance of the optimal airfoils is discussed. Secondly, the convergence efficiency of the six simulation handling strategies is compared and distinctions between the MSES and SU2 optimisation runs are determined. Lastly, the frequency and characterisation of simulation failures in the search space is explored, drawing conclusions on the differences between low- and high-fidelity simulation for supercritical airfoil optimisation.

6.1 Airfoil optimisation results

Due to the small evaluation times of MSES it is possible to obtain an approximated Pareto Front of the trade-off between the weighted lift-to-drag and the drag-divergence Mach number. Multiple surrogate-based optimisation runs are collated together and the combined PF is extracted to initialise the NSGA-II population. This step ensures the approximated PF is not dominated by any of the surrogate-based optimisation runs. The NSGA-II population size is set to 100 and 1,500 generations are evolved, equalling a total of 150,000 function evaluations. The obtained Pareto front and subset of Pareto-optimal airfoils are visualised in Figure 4 – (a). Figure 4 – (b) compares the aerodynamic performance of the obtained optima to the baseline airfoil, SC(2)-0412, for a Mach number sweep at $Re = 30 \cdot 10^6$ and lift coefficient $C_L = 0.4$ using MSES.

The convergence difficulty of the airfoil design problem is moderately high due to the non-linear mapping between objectives and design variables. This aspect is magnified by the high occurrence of simulation failures. The geometric and lift constraints are easy to satisfy and are not a significant challenge for the optimiser. However, the disjointed regions of the PF indicate a high degree of simulation failures. After the initialisation, 96% of airfoil shapes evaluated with MSES fail, which is more severe for the weighted lift-to-drag objective than the drag-divergence Mach number. After 1,500 generations, the simulation failure rate near the vicinity of the PF decreases to 45%. This proportion of simulation failures is significantly higher than those reported in previous studies [1, 2, 3, 14, 15], which highlights the simulation failure handling difficulty of the considered airfoil design problem. Importantly, this enables the robust design of simulation failure handling mechanisms in surrogate-based optimisation, where almost all of the initial population is undefined in objectives and some constraints.



(a) Convergence of NSGA-II for 1,500 generations.

(b) Comparison of wave-drag build up for the four optima and the baseline SC(2)-0412 airfoil.

Figure 4 – Visualisation of optimisation results obtained from NSGA-II and the MSES solver. The colour-map represents generations and the Pareto Front is shown in red with four optimal airfoils annotated in black.

The optimal sections visualised in Figure 4 – (a) show that MSES predictions correctly drive the optimiser towards supercritical profiles. A higher M_{dd} is obtained through a blunter nose and flatter upper section, delaying the onset of strong shock formations, at the expense of higher drag for the operating Mach numbers. At the other extreme of the front, a lower M_{dd} and higher weighted lift-to-drag solution, leads to a sharper undercut leading edge nose and a more rounded upper surface. As a result, a reduction in aft-loading is obtained by decreasing the trailing edge angle γ_{TE} . The optimal solutions feature similar aspects to the conclusions made in the third phase design of the NASA supercritical airfoils [18]. Figure 4 – (b) shows that all optimal airfoils achieve a higher (predicted) drag-divergence Mach number than the baseline SC(2)-0412. Optima 1 and 2 achieve equal or lower drag coefficients at Mach numbers below $M = 0.74$, with reasonably flat drag curves until the critical Mach number is reached. On the contrary, optima 3 and 4 achieve a higher M_{dd} value by maintaining the $\partial C_d / \partial M$ rate-of-change below 0.1 with a more pronounced build up in drag and a higher base-drag. This results in inferior aerodynamic performance below $M = 0.78$. The obtained optimisation results demonstrate the aerodynamic performance of the baseline SC(2)-0412 airfoil can be improved using a low fidelity solver. However the varied wave-drag characteristics observed makes the selection of an optimal design from the Pareto Front a non-trivial task for the designer.

6.2 Simulation failure handling effect on convergence

A convergence study on the six proposed simulation failure handling strategies is conducted using both low- and high-fidelity solvers. Figure 5 – (a) shows the convergence of the Hypervolume Coverage (HVC) for the MSES runs, and Figure 5 – (b) shows the corresponding simulation failure rates during the optimisation history. The equivalent is shown for the high fidelity SU2 optimisation runs in Figure 6 – (a) and (b), respectively. The final median Pareto fronts for each SFH strategy are shown for MSES and SU2 runs in Figure 7 –(a) and Figure 7 – (b), respectively. Table 2 summarises the mean convergence scores and simulation failure ratios for the MSES and SU2 optimisation runs.

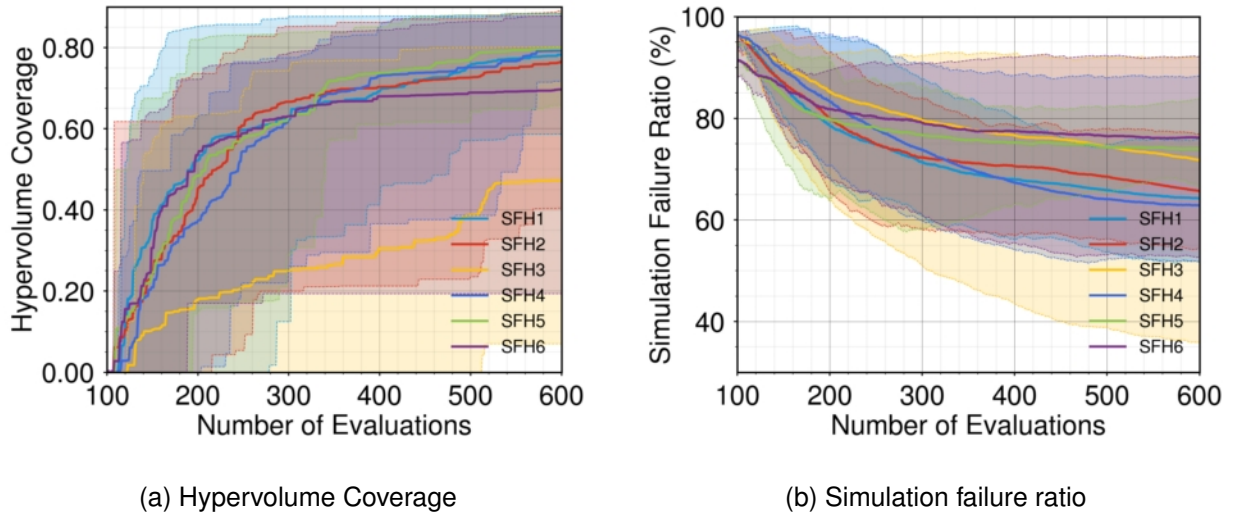


Figure 5 – Convergence and simulation failure ratios for SFH strategies using MSES. Solid lines represent the mean performance and shaded areas are the best and worst performance of 10 independent runs.

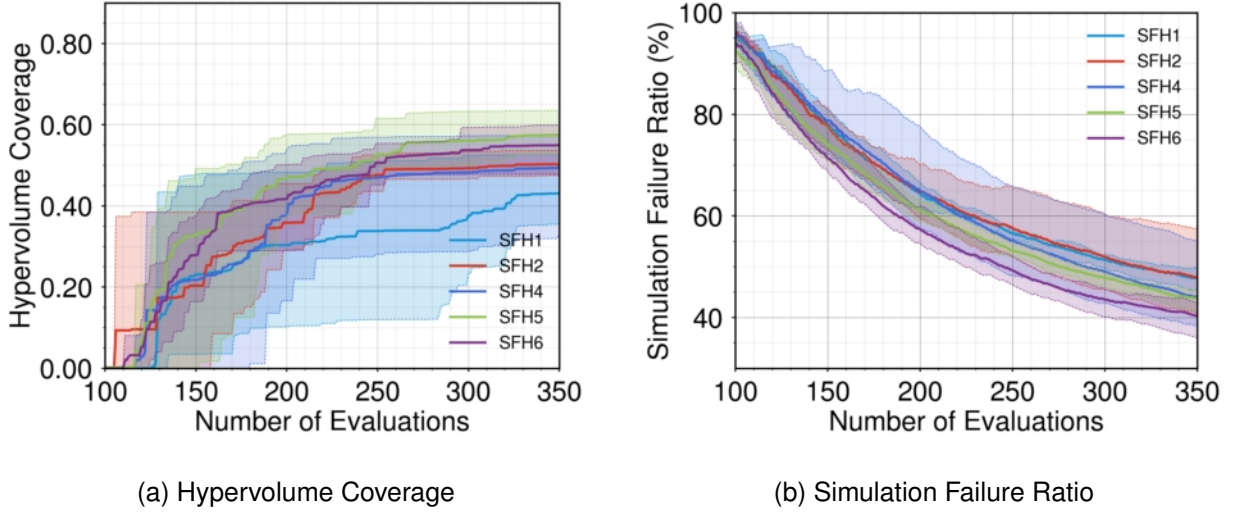


Figure 6 – Convergence and simulation failure ratios for SFH strategies using SU2. Solid lines represent the mean performance and shaded areas are the best and worst performance of 5 independent runs.

The convergence studies highlight a number of differences when conducting optimisation with different solver fidelity. The best mean HVC score of approximately 79% is obtained with strategies SFH4 and SFH5 for the MSES runs, in comparison to 58% for SFH5 in the SU2 runs. The reduction in convergence from the SU2 runs is explained by two aspects; 1. the computational budget is reduced from 50D to 25D due to long run times, and 2. the Pareto front used to calculate the HVC is the MSES approximation of the true front. Validation of the aerodynamic solvers showcased an under prediction of drag for MSES and very similar drag-divergence Mach number predictions for both solvers. This behaviour is seen in Figure 7 – (a) and Figure 7 – (b), where the MSES non-dominated fronts have equivalent M_{dd} objectives to the SU2 runs, but higher lift-to-drag predictions, thus closer to the PF obtained from Figure 4 – (a). Comparing MSES runs in isolation, it is observed that imputing values for failed evaluations (SFH3) does not perform well, with a 59% reduction in HVC relative to the best-performing strategy and no improvement in the final SFR values. For this reason, this strategy is omitted from the computationally-expensive SU2 runs. Penalty methods introduce significant distortion in the regression models and will likely result in poor surrogate modelling accuracy.

Table 2 – Summary of convergence results for the MSES and SU2 optimisation runs of the six SFH strategies.

SFH strategy	mean HVC (%)		mean initial SFR (%)		mean final SFR (%)	
	MSES	SU2	MSES	SU2	MSES	SU2
1	78.3	43.2			65.8	56.7
2	76.5	50.3			68.5	57.6
3	47.3	–	96.3	96.3	74.4	–
4	79.2	49.7			64.1	55.2
5	79.7	57.5			74.3	53.3
6	69.7	54.9	91.6	93.2	76.5	49.3

Exploiting cheap constraint information throughout the entire optimisation process (SFH6) results in pronounced differences between MSES and SU2. For the low-fidelity optimisation runs, the effect of SFH6 is detrimental compared to SFH5, whereas for SU2, strategy SFH6 is closer to SFH5 in performance with the lowest mean SFR of 49%. The use of surrogates for all constraints in SFH5 does not drastically increase the modelling errors for this problem, compared to SFH6, which evaluates the 15 geometric constraints exactly. This suggests that exploiting the constraint surrogate predictions has the benefit of by-passing local optima and feasibility boundaries in the case of MSES, thus achieving

more efficient convergence for the given computational budget. The discrepancy in SFH6 convergence between MSES and SU2 likely relates to the higher simulation failure rates observed by the former, especially near the Pareto front (Section 6.3).

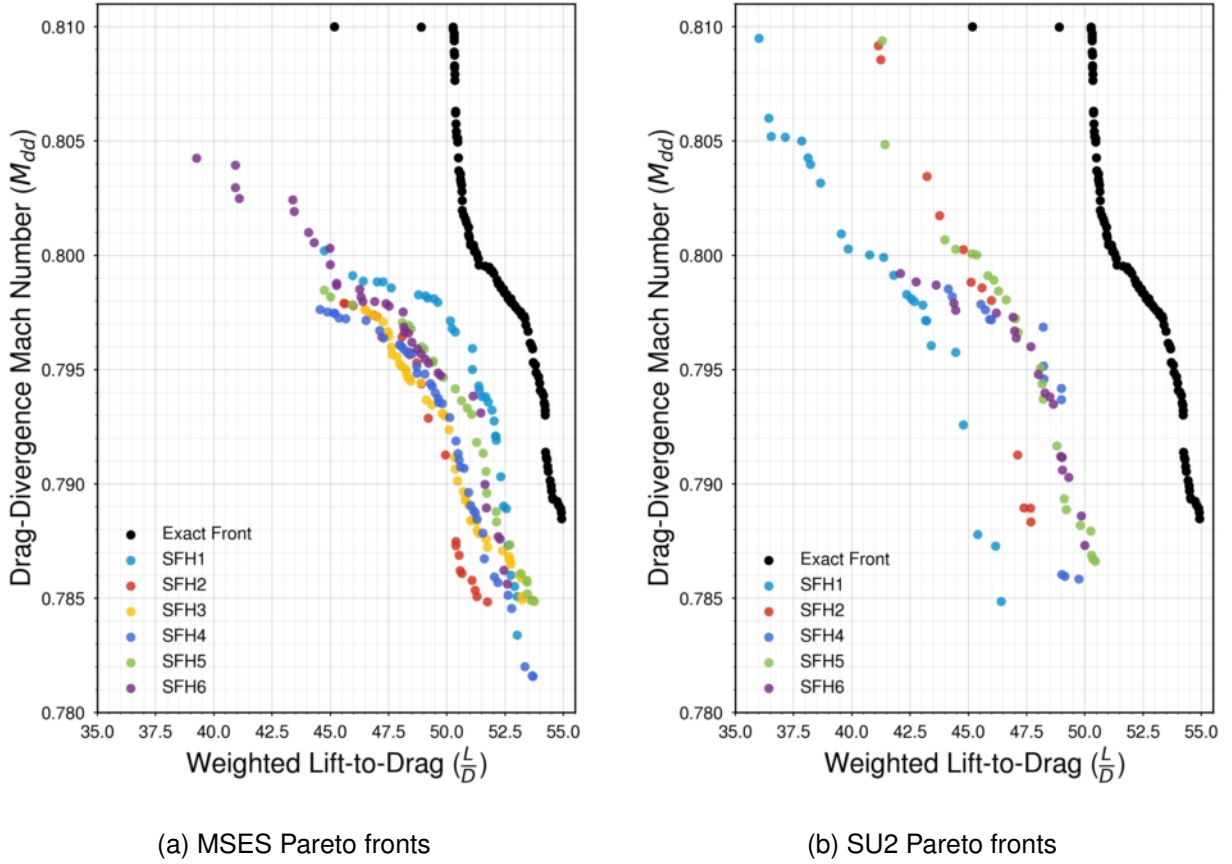


Figure 7 – Final median optimal front for SFH strategies, comparing MSES and SU2 runs.

Strategies SFH1, 2, 4, and 5 obtain near-identical mean HVC scores for MSES, with slight differences in convergence robustness between optimisation runs. SFH5 is the most robust with the lowest variance between the best and worst optimisation runs, however, with a higher final SFR in the population than all strategies besides SFH6. For the SU2 runs, a greater spread in convergence is seen for the above-mentioned strategies. Initialising the population from only geometrically-feasible solutions is most promising when using SU2, as higher HVC scores are achieved compared to strategies SFH1, 2, and 4. The pre-filtering and classification methods in SFH2 and SFH5 are superior to SFH1, which performs the worst for SU2. It appears that simply ignoring failed evaluations is more appropriate when the occurrence and distribution of simulation failures is more prominent and randomised (see Figure 8), which explains why SFH1 performs as well as the other strategies for MSES but not for SU2 runs.

The final SFR scores of the different SFH strategies show that the average failure rate in the final population is 70.6% and 54.4% for MSES and SU2, respectively. As the infill computational budget is halved for the high-fidelity runs, this outcome is significant and highlights that optimisation with SU2 is more robust than MSES for the same high-subsonic flow conditions. It is assumed the higher fidelity modelling of the RANS equations in SU2 accounts for this behaviour. Additionally, the spread in the SFR between optimisation runs is much larger for MSES than for SU2, however this is in part due to only evaluating five instead of ten optimisations runs. Finally, geometric constraint filtering of the initial population only has a small effect on the initial SFR values for both MSES and SU2 runs. This indicates that the initial airfoil geometries generated, although geometrically feasible, produce complex flow behaviour that results in solution divergence or poor convergence, causing both low-

and high-fidelity solvers to fail frequently. Strategies SFH5 and SFH6 reduce the initial simulation failure rate by at most 5% for both solvers, compared to strategies SFH1–4 using a standard LHS initialisation.

Given the obtained convergence results, the following conclusions can thus be derived:

1. Penalty methods must be avoided when training surrogate models and handling simulation failures. They lead to inferior convergence for a surrogate-based optimiser regardless of the solver fidelity.
2. Equivalent optimisation convergence can be obtained with different simulation handling strategies despite different levels of simulation failures. The proportion of simulation failure rates will heavily depend on the problem difficulty and vary drastically between different initialisations. As such, less failed evaluations for a simulation failure handling strategy does not necessarily correspond to improved algorithm convergence and vice versa.
3. The use of cheap constraint information to pre-filter the initial population before expensive evaluation decreases the initial simulation failure rate by approximately 5% both for low- and high-fidelity solvers. This improves the robustness of the convergence for multiple optimisation runs, as seen for strategy SFH5.
4. Exploiting cheap geometric constraint information throughout the entire optimisation process might not yield optimal results depending the simulation failure occurrence and distribution in the design space. For both MSES and SU2 solvers, strategy SFH5 is most appropriate, as the use of surrogates for all constraints enables efficient search and identification of optimal regions in the design space.

6.3 Simulation Failure Characterisation

The occurrence and distribution of simulation failures in the design space is investigated using two high-dimensional data visualisation techniques, namely Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP) [35] and parallel-coordinate plots [36], shown in Figure 8 and Figure 9, respectively. The UMAP visualisations are obtained by projecting the entire optimisation run history, and annotating the points that are simulation failures and geometrically-infeasible.

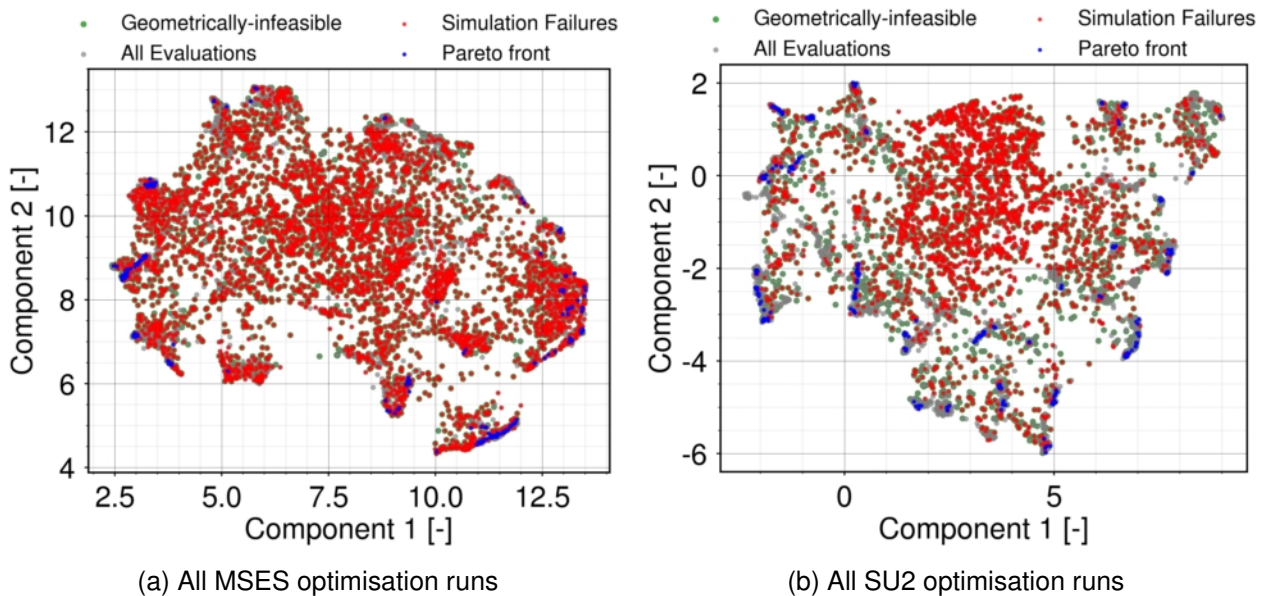


Figure 8 – Projection of all evaluated and simulation failure points for strategy SFH1 using UMAP.

For both MSES and SU2 runs, the non-dominated points are projected at the extremities of the feature map. Notably, the optimal design variables for MSES runs are surrounded by clusters of simulation failures, whereas for SU2, less failures occur closer to the optimal regions. This phenomenon, in conjunction with overall lower simulation failure rates, indicates that SU2 is more robust in transonic regimes for supercritical airfoil design than MSES. Simulation failures for SU2 show greater global structure compared to MSES, however, the frequency and randomness of failures makes it very difficult to predict and mitigate their occurrence during optimisation. Comparing geometrically-infeasible solutions to failed evaluations reveals that 71% of simulation failures are attributed to geometric constraint violations for MSES and 67% for SU2. Although strategy SFH6 shows mixed results in replacing geometric surrogate constraint predictions with exact evaluations, there appears to be a strong correlation between simulation failures and airfoil geometries that violate the 15 geometric constraints imposed of this airfoil problem.

Using parallel coordinate plots, similar observations to the UMAP visualisations can be made. Failures are widespread for all combinations of design variables, including around the non-dominated solutions. For MSES, the clustering of failures is again more prevalent than for SU2, as shown in Figure 8. Whilst UMAP provides a visualisation of the design space, Figure 9 enables direct comparison of the optimal design variables. Both MSES and SU2 solvers guide the optimiser towards very similar geometries, besides the Bezier nodes X_4 and X_8 , which show the greatest variation. These control points affect the upper leading edge and lower surface maximum thickness regions, respectively. It should be noted that the differences in PF coverage for the low- and high-fidelity runs explains some of the variance observed. As shown, the occurrence and distribution of simulation failures in the design space is non-trivial. SU2 appears to be more robust than MSES once the optimal region of the design space is identified by the surrogate-based optimiser. This is seen by the reduced number of failed evaluations surrounding the Pareto front, which also results in greater variation in the optimum design variables. Finally, a strong correlation between geometric constraint violations and simulation failures is again observed, thus the key to reducing wasted computational effort likely lies in filtering out geometrically-infeasible solutions before expensive evaluation and setting up appropriate constraints during problem design.

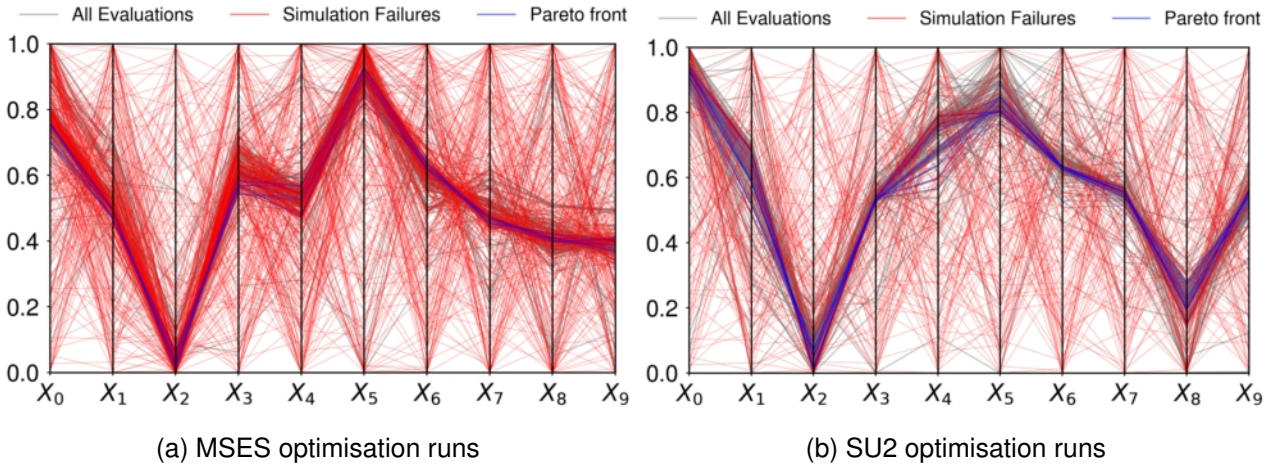


Figure 9 – Projection of all evaluated and simulation failure points for strategy SFH1 using parallel coordinates.

7. Conclusion

Surrogate-based shape optimisation has been applied to the design of a supercritical airfoil subject to a high degree of simulation failures using low- and high-fidelity solvers MSES and SU2, respectively. Convergence studies on six simulation failure handling strategies, inspired from literature, were conducted to determine the occurrence and distribution of failed evaluations. Furthermore, two high-dimensional data visualisation tools were used to analyse the differences in failures between solvers of varying fidelity and the resulting optimal design variables of the Pareto fronts obtained.

Results show that the frequency of simulation failures is approximately 23% lower for SU2 compared to MSES at half the infill computational budget, and optimal regions of the design space are much less failure-prone for SU2 than for MSES. Moreover, a strong correlation between geometric constraint violation and simulation failures of 71% and 67% is determined for MSES and SU2, respectively. From the six strategies investigated, classical penalty methods are found to not be appropriate for handling simulation failures, due to the impact on surrogate modelling accuracy. No significant differences are found, besides robustness between the best and worst optimisation runs, for failed evaluation pre-filtering and classification in simulation failure handling strategies (SFH) two and four. A simple strategy that ignores failed evaluations (SFH1) is appropriate when the occurrence of simulation failures is high and very randomised, as with MSES. However, when there is greater global structure to simulation failures in the design space, SFH1 performs poorly compared to the other tested strategies, as shown for SU2 results. Pre-filtering the initialisation population for geometric feasibility appears to reduce the number of simulation failures by approximately 5% initially. Nevertheless, the strong correlation between the geometric constraint violation and simulation failures indicates that this information, if available to the designer, should be exploited when conducting aerodynamic shape optimisation. Lastly, it is expected that the outcomes observed in this study are highly-problem dependent, and also reflect an extreme, yet common, occurrence of simulation failures. Further work is warranted to capture the general trends in simulation failure handling for surrogate-based optimisation on a wide range of (aerodynamic) shape optimisation problems subject to simulation failures.

8. Contact Author Email Address

For further communication, please contact the corresponding author's email:
 Juan Rasines Mazo, <mailto:juan.rasinesmazo@sydney.edu.au>

9. Copyright Statement

The authors confirm that they, and/or their company or organization, hold copyright on all of the original material included in this paper. The authors also confirm that they have obtained permission, from the copyright holder of any third party material included in this paper, to publish it as part of their paper. The authors confirm that they give permission, or have obtained permission from the copyright holder of this paper, for the publication and distribution of this paper as part of the ICAS proceedings or as individual off-prints from the proceedings.

References

- [1] Arsenyev I., Duddeck F., and Fischersworring-Bunk A. Adaptive surrogate-based multi-disciplinary optimization for vane clusters. In *Turbo Expo: Power for Land, Sea, and Air*, volume 56659, page V02CT45A002. American Society of Mechanical Engineers, 2015.
- [2] Baert L., Dumeunier C., Leborgne M., Sainvitu C., and Lepot I. Agile sbo framework exploiting multisimulation data: Optimising efficiency and stall margin of a transonic compressor. In *Turbo Expo: Power for Land, Sea, and Air*, volume 51029, page V02DT46A019. American Society of Mechanical Engineers, 2018.
- [3] Wilke G. Variable-fidelity methodology for the aerodynamic optimization of helicopter rotors. *AIAA Journal*, 57(8):3145–3158, 2019.
- [4] He Y., Sun J., Song P., and Wang X. Variable-fidelity expected improvement based efficient global optimization of expensive problems in presence of simulation failures and its parallelization. *Aerospace Science and Technology*, 111:106572, 2021.
- [5] Andersson B., Andersson R., Håkansson L., Mortensen M., Sudiyo R., and Van Wachem B. *Computational fluid dynamics for engineers*. Cambridge university press, 2011.
- [6] Palar P. S., Liem R. P., Zuhail L. R., and Shimoyama K. On the use of surrogate models in engineering design optimization and exploration: The key issues. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pages 1592–1602, 2019.
- [7] Yondo R., Bobrowski K., Andrés E., and Valero E. A review of surrogate modeling techniques for aerodynamic analysis and optimization: current limitations and future challenges in industry. *Advances in evolutionary and deterministic methods for design, optimization and control in engineering and sciences*, pages 19–33, 2019.
- [8] Shahjahan S., Gong A., Moore A., and Verstraete D. Optimisation of proprotors for tilt-wing evtol aircraft. *Aerospace Science and Technology*, 144:108835, 2024.
- [9] Jin Y. Surrogate-assisted evolutionary computation: Recent advances and future challenges. *Swarm and Evolutionary Computation*, 1(2):61–70, 2011.
- [10] MacNeill R. and Verstraete D. Propeller optimisation for an electrically-powered tactical uas. In *Proceedings of the 31st Congress of the International Council of the Aeronautical Sciences, Belo Horizonte, Brasil*, 2018.
- [11] Verstraete D. Multi-disciplinary optimisation of medium altitude long endurance uavs. In *31st International Congress of the Aeronautical Science, ICAS-2018-416*, 2018.
- [12] He C., Zhang Y., Gong D., and Ji X. A review of surrogate-assisted evolutionary algorithms for expensive optimization problems. *Expert Systems with Applications*, page 119495, 2023.
- [13] Adler E. J., Gray A. C., and Martins J. To cfd or not to cfd? comparing rans and viscous panel methods for airfoil shape optimization. In *33rd Congress of the International Council of the Aeronautical Sciences*, 2022.
- [14] Tenne Y. and Armfield S. A versatile surrogate-assisted memetic algorithm for optimization of computationally expensive functions and its engineering applications. *Success in evolutionary computation*, pages 43–72, 2008.
- [15] Sainvitu C., Iliopoulou V., and Lepot I. Global optimization with expensive functions-sample turbomachinery design application. In *Recent Advances in Optimization and its Applications in Engineering: The 14th Belgian-French-German Conference on Optimization*, pages 499–509. Springer, 2010.

- [16] Drela M. A User's Guide to MSES 3.05. *Massachusetts Institute of Technology (MIT), Cambridge*, 2007.
- [17] Economon T. D., Palacios F., Copeland S. R., Lukaczyk T. W., and Alonso J. J. Su2: An open-source suite for multiphysics simulation and design. *Aiaa Journal*, 54(3):828–846, 2016.
- [18] Harris C. D. Nasa supercritical airfoils: A matrix of family-related airfoils. Technical report, 1990.
- [19] Shikhar Jaiswal A. Shape parameterization of airfoil shapes using bezier curves. In *Innovative Design and Development Practices in Aerospace and Automotive Engineering: I-DAD, February 22-24, 2016*, pages 79–85. Springer, 2016.
- [20] Spaid F. W. and Hakkinen R. J. On the boundary layer displacement effect near the trailing edge of an aft-loaded airfoil. *Zeitschrift für angewandte Mathematik und Physik ZAMP*, 28:941–950, 1977.
- [21] Drela M. Xfoil: An analysis and design system for low reynolds number airfoils. In *Low Reynolds Number Aerodynamics: Proceedings of the Conference Notre Dame, Indiana, USA, 5–7 June 1989*, pages 1–12. Springer, 1989.
- [22] Smith M. J., Potsdam M., Wong T.-C., Baeder J. D., and Phanse S. Evaluation of computational fluid dynamics to determine two-dimensional airfoil characteristics for rotorcraft applications. *Journal of the American Helicopter Society*, 51(1):70–79, 2006.
- [23] Koss D., Steinbuch M., and Shepshelovich M. Design and experimental evaluation of a high-lift, mild-stall airfoil. In *12th AIAA Applied Aerodynamics Conference*, 1994.
- [24] Liebeck R. H. Low reynolds number airfoil design for subsonic compressible flow. In *Low Reynolds Number Aerodynamics: Proceedings of the Conference Notre Dame, Indiana, USA, 5–7 June 1989*, pages 314–330. Springer, 1989.
- [25] Fujino M., Yoshizaki Y., and Kawamura Y. Natural-laminar-flow airfoil development for a lightweight business jet. *Journal of Aircraft*, 40(4):609–615, 2003.
- [26] Morgado J., Vizinho R., Silvestre M., and Páscoa J. Xfoil vs cfd performance predictions for high lift low reynolds number airfoils. *Aerospace Science and Technology*, 52:207–214, 2016.
- [27] Carreño Ruiz M., Renzulli L., and D'Ambrosio D. Airfoil optimization for rotors operating in the ultra-low reynolds number regime. *Physics of Fluids*, 35(10), 2023.
- [28] Cakmakcioglu S. C., Bas O., Mura R., and Kaynak U. A revised one-equation transitional model for external aerodynamics. In *AIAA Aviation 2020 Forum*, page 2706, 2020.
- [29] D P. Construct2d: Computational fluid dynamics structured grid creator for 2d airfoils, 2013. Available at <https://sourceforge.net/projects/construct2d/>.
- [30] Shahjahan S., Emmerson B., and Verstraete D. On the choice of turbulence model for the simulation of airfoils at reynolds numbers below 200,000. In *34th International Congress of the Aeronautical Science, ICAS-2024-327*, 2024.
- [31] Harris C. D. Aerodynamic characteristics of the 10-percent-thick nasa supercritical airfoil 33 designed for a normal-force coefficient of 0.7. Technical report, 1975.
- [32] Deb K., Pratap A., Agarwal S., and Meyarivan T. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation*, 6(2):182–197, 2002.
- [33] Bagheri S., Konen W., and Bäck T. Comparing kriging and radial basis function surrogates. In *Proc. 27. Workshop Computational Intelligence*, pages 243–259. Universitätsverlag Karlsruhe, 2017.

- [34] Huband S., Hingston P., While L., and Barone L. An evolution strategy with probabilistic mutation for multi-objective optimisation. In *The 2003 Congress on Evolutionary Computation, 2003. CEC'03.*, volume 4, pages 2284–2291. IEEE, 2003.
- [35] McInnes L., Healy J., and Melville J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.
- [36] Moustafa R. E. Parallel coordinate and parallel coordinate density plots. *Wiley Interdisciplinary Reviews: Computational Statistics*, 3(2):134–148, 2011.