
Peer reviewed version

Link to published version (if available):
10.2514/6.2016-1304

Link to publication record in Explore Bristol Research
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via AIAA at http://arc.aiaa.org/doi/abs/10.2514/6.2016-1304. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research
General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/pure/about/ebr-terms
Optimization of Vane-Type Vortex Generators for Tiltrotor Wings using Computational Fluid Dynamics

R.L.T. Bevan*, D.J. Poole†, C.B. Allen‡, T.C.S. Rendall§

Department of Aerospace Engineering, University of Bristol, Bristol, BS8 1TR, U.K.

Tiltrotor wings are thick and highly loaded airfoils, and are thus particularly susceptible to separation and early onset buffet and vortex generators are commonly used to alleviate these issues. Hence, the design of counter-rotating vortex generators on tiltrotor wings to control separation is considered. A representative tiltrotor wing airfoil (a modified NACA 64(4)-421) has been tested at a Reynolds number of 7.5 million ($M = 0.24$) and $C_L = 1.3$ to demonstrate that a large separation is problematic in this flow case. Five parameters have been introduced to quantify the vortex generator design space (length, aspect ratio, angle, chord location and spacings ratio) and Latin Hypercube sampling of this design space has been undertaken using high fidelity CFD to construct a radial basis function based surrogate model. This model has been interrogated and refined using an efficient adaptive sampling method which incorporates both space-filling and local refinement properties in order to achieve an optimum solution. To determine this optimum, a global search algorithm based on differential evolution was used for optimization of the model. The results obtained from the optimization eliminated the separation experienced by the clean geometry (from 11.1% to 0%), whilst reducing the drag obtained from the best Latin Hypercube sample by 7%.

I. Introduction

Tiltrotor aircraft combine the vertical lift capability of a helicopter with the speed and endurance of a conventional fixed-wing aircraft. In forward flight, the aircraft lift is provided by the wings. However, due primarily to structural design considerations, these are very stiff, low-aspect ratio wings, which inhibit efficient aerodynamic performance. Due to a short span and a typical cruise Mach between 0.4 and 0.5, these wings experience a high loading; a typical cruise lift coefficient for a tiltrotor being above 1.0. Structural requirements also result in a thick aerofoil. The Bell XV-15 utilised a NACA 64A223 in this regard. These design factors therefore make tiltrotor aircraft particularly susceptible to stall and early onset buffet, so boundary layer control is critical and upper surface flow control devices are an important consideration.

Flow control devices are adopted regularly in aerodynamic design, to control boundary layer behaviour, and are often retro-fitted where problems are identified. Occasionally, complex active systems have been implemented to control boundary layer growth, including boundary layer suction, blowing, and bleeding, but by far the most common methods are passive schemes, with vortex generators (VGs) the most conventional approach. These involve simple small flat plates normal to the surface, usually rectangular in shape, inclined to the freestream flow. VGs were utilised on the Bell XV-15 tiltrotor to eliminate premature stall and retain attached flow up to an angle of attack of 15 degrees. The V-22 Osprey tiltrotor also employs VGs; at mid to high angles of attack, experimental data shows that wing-mounted VGs suppress the early onset of separation. The majority of VGs protrude into the external flow, i.e. above the boundary layer, but there have also been immersed, or sub-layer, VGs adopted. The role of a VG is to induce vortical flow via the flow separation from its sharp upper edge. The vortex which subsequently propagates downstream entrains

---

*Researcher. Email: r.bevan@bristol.ac.uk
†Researcher. Email: d.j.poole@bristol.ac.uk
‡Professor of Computational Aerodynamics. Email: c.b.allen@bristol.ac.uk
§Lecturer. Email: thomas.rendall@bristol.ac.uk
higher energy, higher momentum flow from outside the boundary layer into the lower energy boundary layer. This re-energises the boundary layer, suppressing separation.

It is necessary to understand the physics that the VGs are introducing into the flow to be able to design them to mitigate against unwanted flow phenomena (e.g. separation), and this is particularly important on tiltrotor wings due to the susceptibility to separation or buffet. To understand and quantify the effects of VGs, computational fluids dynamics (CFD) approaches are often employed. However, the modelling of VGs by conventional CFD approaches poses difficult issues in capturing the vortex and its convection downstream. This often requires fine numerical meshes on geometries where the size of the flow control device is orders of magnitude smaller than the global object. Furthermore, a sufficiently high fidelity numerical scheme must be selected, often requiring Reynolds’s averaged Navier-Stokes (RANS) solutions with turbulence models.

In order to determine an optimized VG geometry, the high fidelity simulation approach can either be coupled directly to an optimization routine or alternatively, it can be incorporated within a surrogate-based analysis. The latter approach uses a limited number of high fidelity runs to train a surrogate model. This model then provides interpolation of the known sample points to obtain a continuous design space approximation away from the data sites. Using this surrogate-based approach allows the relatively cheap design of the VG compared to the directly coupled approach. Design space interrogation methods can then locate an optimum VG design without the need for large numbers of simulations. This surrogate modelling approach for design of VGs has received only limited coverage in previous publications, including by the use of lower-fidelity simulation approaches.

The surrogate-based approach relies on limiting the number high fidelity sample points within the design space, and thus efficient design space interrogation methods are desired. Two sampling strategies for constructing surrogate models were examined in Mackman et al. for construction of aerodynamic models. The two methods employed were adaptive, i.e. information from successive iterations is used to reconstruct the surrogate model between iterations. It was determined that both methods were suitable for aerodynamic application, and both provided better than traditional one-stage or sequential space filling algorithms. An alternative adaptive sampling method for surrogate-based modelling of multidimensional design space, combines both space-filling properties and local refinement.

A surrogate modelling approach was applied to tilt-rotor aircraft wings in Bevan et al. This earlier work introduced a framework of statistical sampling and surrogate modelling, together with a preliminary examination to highlight the potential of the approach. This initial examination utilised a simplified design space, which was optimised for a single design point. The work presented in this paper refines the prior framework and broadens the design space through both increasing variable range and also improving and expanding on the parameters considered. Finally, adaptive sampling is introduced in order to augment interrogation of the surrogate model and further optimise the resulting VG configuration.

II. Computational Modelling of Vortex Generators

The physics of a flowfield when a vortex generator (VG) is placed into the flow differs considerably from a clean geometry. The VG produces a vortex which encourages mixing of turbulent flow downstream of the device, often used to avoid unwanted separation. The exact physics of flow control using VGs is complicated and often requires fine computational meshes with high fidelity flow simulations. The detail in this section gives examples of the type of simulations used for VG modelling.

II.A. Mesh Generation

The precise physical effects and the topology of VGs makes for difficult modelling, both in set-up and in solution methods. Multi-block structured meshes are difficult to produce for an aerofoil with a VG due to the large differences in scale that exist between the parent aerofoil and the flow control device, and the sudden changes in curvature when progressing over the surface of the aerofoil. Furthermore, to effectively capture and convect downstream the vortex produced by the device, fine resolution grids are often required around the VG and in its wake, resulting in the exponential growth of mesh sizes when compared to a clean wing counterpart. Despite the difficulties in producing structured meshes, this meshing approach is generally considered to produce the highest quality meshes that are most suitable to model flow around VGs and so many examples can be found.

To alleviate some of the issues with structured mesh generation, unstructured meshes can instead be
used. The advantage of unstructured meshes is that the algorithms for producing these meshes typically use space-filling or marching approaches and so only require surface and farfield geometry, making their production considerably simpler than multiblock structured meshes. However, to capture the physics of the VG in the flow, a structured prismatic layer must exist on the surface. Examples include.

II.B. Flow Solver

The suitable selection of a CFD solver is also imperative to ensure the complete convection of the vortex downstream from its source without too much dissipation being added. The lowest fidelity simulation of VGs that has been done is by Kerho and Kramer who used XFOIL with a modification to the stress-transport equation at the VG location to mimic the effect of a VG. This was demonstrated to reduce the drag values near the separated region of an aerofoil.

In general, full flowfield analysis of a three-dimensional geometry must be performed to fully capture the physics of the vortex generator in the flow, and this must be done by high fidelity RANS, or even LES and DNS solutions. If using a RANS simulation approach then it is of paramount importance to select an appropriate turbulence model to ensure detailed capture of the vortex. Two common turbulence models are Spalart-Almaras (SA) and k-ω SST which have both been used to model flow around a VG. Both tend to under-predict streamwise vorticity of the vortex, though the SST model generally matches the experiment better than SA. This can be alleviated somewhat by including the strain-vorticity modification to the SA model. Yaras and Grosvenor also compared both models with the k-ε and standard k-ω, and showed that both SA and SST outperformed the the k-ε and k-ω models, but that SA perhaps demonstrated the best balance between prediction, robustness and computational efficiency. The work of Tai utilised a modified Baldwin-Lomax algebraic model, as well as Menter’s SST model as a supplement to ensure the physical flow features were captured, on a V-22 Osprey simulation with VGs.

A lower-fidelity approach using embedded models within a CFD framework is also a popular approach. Models such as the vortex-source and lifting-force have been shown to produce results that mimic the effect of a vortex generator placed in the flow without the need to mesh and solve the full geometry.

To fully capture all the flow-physics associated with the VG, high-fidelity, real-time, simulations via large eddy simulation (LES) or direct numerical simulation (DNS) methods are required, though these are prohibitively expensive, particularly for design. As such, historically it has been difficult to perform such simulations for large Reynolds numbers, however, with the continual increase in computational resources available these are becoming more common. The Lattice-Boltzmann method has also been used to produce time-dependent flow simulations around VGs.

III. Simulation Approach

As outlined above, the modelling of the vortex generator and resulting flow is not a trivial task; the suitable capture of the physics requires high fidelity solvers and high density numerical meshes. The primary purpose of the research presented in this paper is the simulation of numerous VG configurations, in order to assess the optimal design considerations. A solution method is firstly developed and presented in this section. This involves obtaining high-fidelity viscous flows using OpenFOAM with a bespoke mesh generation tool to produce high quality structured viscous meshes around a representative aerofoil section with a vortex generator.

III.A. Mesh Generation

All the meshes generated utilised the C-grid approach. A representative tiltrotor aerofoil, which is a modified NACA 64(4)-421 section was used, with a 0.4% chord blunt trailing edge. A two-dimensional four-block structured multiblock mesh was generated initially, and three views are presented in figure. This is for a mesh with 769 points on the aerofoil surface, 97 streamwise points on either side of the wake line, 65 points across the blunt trailing edge, and 129 points in the normal direction.

To include the vortex generator, the two-dimensional mesh is extruded to the correct part-span position, and a high quality mesh deformation scheme used, to rotate the mesh to the required vortex generator.

*http://www.OpenFOAM.org/
incidence, and pull in mesh cells from upstream, downstream, above and either side of the generator to match the required local spacings. This is achieved via several local radial basis function deformations.

Figure 2 shows the resulting wing surface mesh near the leading edge and the surface with starboard plane, together with a closer view of the VG; this is for a sample point with 9.85% chord wide surface, with rectangular vortex generator $1.88\% \times 1.31\%$ chord length and height, with leading edge at 27.83% chord from leading edge, 0.5% chord from port plane, at 23.75 degrees incidence to the freestream. The number of spanwise planes extruded in dependent upon the spanwise domain width required. In this instance, 127 spanwise planes were extruded.

The final stage in the mesh generation process reduces the cell count away from sections of interest. This is undertaken through point removal and combining of hexahedra cells to form polyhedra. The mesh reduction process typically involves a 50-60% cell count reduction.

III.B. CFD Solver

Once a suitable mesh has been produced, this is used to model the viscous flow around the wing with VGs using OpenFOAM with a RANS model. OpenFOAM is an open source CFD software package, capable of solving problems such as: complex fluid flows involving chemical reactions, turbulence and heat transfer, solid dynamics and electromagnetics.

The flow simulations utilise a unstructured solver based upon the SIMPLE algorithm (Semi-Implicit Method for Pressure-Linked Equations), an iterative procedure to solve the Navier-Stokes equations for steady-state problems. This iterative procedure, as originally implemented in OpenFOAM, relies on basic residual checking for steady state convergence criteria. After modification, steady state convergence is assessed using the standard deviation of the force coefficients.

OpenFOAM has numerous turbulence models that can be utilised, although these models are not necessarily considered standard implementations (e.g. SA-fv3). After modifications, this work utilises the standard Spalart-Allmaras (SA) model as defined by NASA Langley’s Turbulence Modelling Resource\textsuperscript{b}. The standard SA model implemented in this work uses the minimum limiter of $0.3 \times \Omega$ for $\tilde{S}$.

\textsuperscript{b}http://turbmodels.larc.nasa.gov
III.C. Example Solutions

The simulations were undertaken at a Reynolds number of 7.5 million, and a Mach number of 0.24. If present, the VG was located at 11.25%. The distance between two counter-rotating pairs was 5.6% and the intra-pair spacing was 0.5%. The VG height and aspect ratio (defined as length/height) were 1.16% and 1.78 respectively, while the VG angle relative to freestream was 20.25 degrees. All percentages are relative to the airfoil chord. No-slip conditions were applied to both the aerofl and the VG. The farfield was located at 100 chord lengths from the aerofl. At the farfield, freestream boundary conditions were imposed, together with a constant pressure at the downstream farfield. While the clean geometry (modelled using a two dimensional mesh) utilises the two dimensional specific span-wise boundary condition within OpenFOAM, the three dimension aerofl and VG utilised slip conditions. Earlier work\textsuperscript{13} determined that this span-wise boundary condition was suitable.

At a $C_L$ of 1.3 (9 degrees angle of attack for aerofl with VG, 10 degrees for clean), the clean aerofl has a significant separated region, as shown in figure 3a.

![Clean Geometry. Streamlines](image)

![Example VG. $C_p$ Span Averaged](image)

![Example VG. $C_f$ Span Averaged](image)

Figure 3: Example solutions around modified NACA 64(4)-421

In contrast, the aerofl with VG does not have a separation region. This separation is assessed by examination of the $x$-direction friction coefficient (with $x$ being the stream-wise direction), which is given in figure 3b. From the figure, it is clear that the separation has been eliminated through introduction of the VG, while the clean geometry has separation after 88% chord. The improved performance relative to the clean geometry is also evident in the plot of the pressure coefficient (figure 3c), with the span-wise $C_p$ variation centred around the immediate downstream of the VG location and then diminishing.

IV. Vortex Generator Design Space

A high-fidelity simulation approach that utilises structured numerical meshes and viscous flow solutions by OpenFOAM has been outlined and tested above for the simulation of flows around VGs on infinite wings. The work in this paper is focussed on the process of performing effective design of the vortex generator (in terms of its planform and location on the wing) and this process is outlined in this section. A full design space of a number of design parameters is considered and sampled by the random Latin hypercube statistical sampling method. These sample points are then simulated using the tools developed to allow investigation of the effect of the design parameters on the flow solution around the VGs.

IV.A. Design Parameters

Five design parameters are investigated in this work. These are VG length ($l$), VG aspect ratio ($AR$), VG setting angle relative to freestream ($\theta$), VG chordwise location ($c$) and the spacing ratio ($s$). The VG aspect ratio is defined as $AR = l/h$, and the spacing ratio is defined as the ratio between the intra-pair spacing ($x_1$) and the inter-pair spacing ($x_2$), where the intra-pair spacing is fixed ($x_1 = 0.5\%$). The use of VG aspect ratio, rather than VG height directly was based upon the work of Bevan et al\textsuperscript{13}. This earlier initial investigation incorporated both VG height and length as individual parameters. Results obtained during the study demonstrated that an $AR < 1$ produced undesirable flow with significant drag penalties. The five
parameters are given in figure 4.

Figure 4: VG Configuration

Rectangular VGs are modelled in this work. The VGs are counter-rotating in nature. The five design parameters were chosen to give a sufficient variation in the flow nature and final force coefficients to design for a chosen objective function. The maximum and minimum values of the design parameters are shown in table 1 and chosen to represent sensible coverage without including extremes of the design space. All parameters are scaled such that the design space is $[0, 1]^5$.

Table 1: Design space range (% based on airfoil chord)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h$</td>
<td>0.2%</td>
<td>2%</td>
</tr>
<tr>
<td>$AR$</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>$\theta$</td>
<td>5°</td>
<td>30°</td>
</tr>
<tr>
<td>$c$</td>
<td>5%</td>
<td>50%</td>
</tr>
<tr>
<td>$s$</td>
<td>5</td>
<td>20</td>
</tr>
</tbody>
</table>

For tiltrotor aircraft wings, which are thick and highly loaded, early onset buffet can be particularly problematic and is typically linked to separation. For this reason, buffet is one of the primary design considerations in this work, and particularly separation at a chosen design condition. The loading on the wings in cruise can be as high as $C_L = 1.0$, however, it is unlikely that buffet in cruise is a problem. The chosen design loading is therefore higher, and set to be $C_L = 1.3$, which is realistic for a manoeuvre loading.

To test for possible buffet, a simulation around the clean aerofoil has been performed for this loading at a representative flight condition ($Re = 7.5 \times 10^6$, $M_\infty = 0.24$, $C_L = 1.3$) and the flowfield is shown in figure 3a. As can be seen, separation at this flight condition is considerable and therefore it is probable that buffet may also be an issue. The vortex generators are used to alleviate the buffet, which is assumed to be caused by a large separation region. Therefore the objective function (equation 1), against which to design the vortex generators, is to minimize separation subject to minimizing the drag penalty. This objective function will balance drag induced by introduction of the VG and that produced via separation.

$$J = C_D + kS$$

$J$ is the objective function to be minimized, $S$ is the span-average separation chordwise location (as a percentage), and $k$ is a constant used to balance the effect of the two. To ensure equal magnitude of both parameters, $k$ is taken to be 0.0005 such that 1% separation equates to adding 5 drag counts to the problem.
IV.B. Latin Hypercube Sampling

To efficiently sample the design space, and reduce the number of function evaluations required from that of brute force sampling, a Latin Hypercube (LHC) statistical sampling approach is used. A Latin Hypercube sampling approach is a statistics-based approach designed to produce an optimum configuration of sample points through the design space; the design space produced has each division in all dimensions sampled by exactly one sample point. Corners of the sample space are added to the sampling as this is important for interpolation, which is the basis of a surrogate modelling approach developed below. Whilst a random LHC was utilised in an optimised LHC offers an improvement in providing a more evenly sampled design space for construction of the surrogate model. Numerous methodologies exist for construction of an optimised LHC, however the detailing of these are beyond the scope of this paper. The approach chosen utilised the Morris and Mitchell\textsuperscript{43} $\phi_p$ criteria

$$\phi_p(X) = \left( \sum_{j=1}^{m} g_j d_j^{-p} \right)^{1/p} \quad (2)$$

where $d$ represents the distance between two points in hypercube $X$ and $g$ represents the number of occurrences of this distance within the hypercube. $d$ can be calculated either with the absolute norm or the Euclidean norm. $p$ is a positive integer. Minimizing $\phi_p$ improves the space-filling properties of the hypercube. As recommended by Morris and Mitchell, this work minimizes $\phi_p$ for $p = [1, 2, 5, 10, 20, 50, 100]$ then the best of the resulting hypercubes is used. Since the possible combinations for a 5 dimension hypercube with only 100 sample points is of the order of $10^{790}$, an optimization algorithm is required. This work uses the Enhanced Stochastic Evolution algorithm proposed by Jin \textit{et al}.\textsuperscript{44} An example of the resulting Latin Hypercube sample is shown in figure 5 for 100 sample points in two-dimensions plus the corner points.

![Figure 5: Latin Hypercube sampled design space (plus corner points)](image)

The Latin Hypercube used in this work consisted of 181 sample and corner points.

IV.C. Results

At all Latin Hypercube samples, a complete polar was run and interpolation used to obtain the angle, and therefore drag, where the lift was the design lift coefficient ($C_L = 1.3$). A brief statistical summary of the results is shown in tables 2 and 3.

From table 2 the standard deviation relative to the mean is significant for both the separation ($S$) and drag ($C_D$), although the objective function $J$ has a marginally smaller standard deviation amongst the points sampled. It is evident from table 2 that considering drag in isolation is not sufficient to reduce (and eliminate) separation. Of the 181 samples, 24.5% have $S < 0.1\%$, but only two (4.4%) of this reduced sample have a $C_D < 0.0210$. These two samples are spread throughout the design space, such that the distance
between these points with the normalised \([0, 1]^5\) design space is 0.85 which equates to 38% of the design space range.

V. VG Design Space Modelling

A full design space investigation of a number of VG design parameter s has been considered above using high-fidelity viscous analysis. A Latin Hypercube sampled design space of the VG geometry and location on the aerofoil has been performed. The next stage of this work is the design optimization of the VG based on the design parameters that have been identified as being of significance. This is achieved by representing the remainder of the design space (i.e. away from the already simulated design points) by an interpolation-based surrogate model that is built around radial basis functions. This interpolation recovers the original data and therefore preserves the high-fidelity nature of the solutions while simultaneously distributing that information throughout the design space.

V.A. Surrogate Modelling

A surrogate model is an interpolation or approximation made from a small sample of data sites that give the value of the design space at those known data sites. The model can be formulated in a number of different ways, though the most common methods are by a polynomial, radial basis functions (RBFs) or Kriging. A brief outline of the three methods is presented below, but reviews of surrogate modelling and surrogate-based optimization have been presented\(^{45–47}\) which the reader is guided to for more in-depth discussions on the formulations and common uses for each approach.

The polynomial method\(^{48}\) involves fitting an order \(m\) polynomial through the given data, often by a least squares fit. The resulting model is not guaranteed to recover any of the existing data, meaning it tends to be more suited to fitting noisy data, where an underlying trend is the desired outcome instead of an exact
model. An RBF model is a linear combination of a number of functions that depend on the radial distance from known data sites.\textsuperscript{43, 44} Unlike polynomial interpolation, RBF interpolation does provide exact recovery at the known data sites, though interpolation of noisy data using RBFs can also be done using a regularization parameter.\textsuperscript{51} Finally, the Kriging approach, pioneered by Krige\textsuperscript{23} and later used for surrogate modelling by Sacks \textit{et al.}\textsuperscript{33} and Jones \textit{et al.}\textsuperscript{34} is a special form of an RBF interpolation that uses a generalised Gaussian equation as the basis function. The function allows variation in the response from variables to variable, and is more statistically-based than a pure Gaussian function.

The surrogate model used in this work is developed around a multivariate interpolation using RBFs. This interpolation method has the advantage of providing exact recovery of data at the known sites, thus preserving the CFD data at the Latin Hypercube sampled points.

The general theory of RBFs is presented by Buhmann,\textsuperscript{25} and Wendland,\textsuperscript{26} and the basis of the method used here is described in detail by Rendall and Aller,\textsuperscript{40} and surrogate specific detail in Mackman \textit{et al.}\textsuperscript{11} Let $f(x)$ be the original function to be modelled, and $f_i$ be the scalar values at $n$ discrete points $x_i, i = 1, ..., n$, where $x_i$ is the vector of inputs at the $i$th sample point in $d$-dimensional space $x^j, j = 1, ..., d$. The set of data points $X = \{x_1, ..., x_n\}$ is confined to a domain $\Omega$ in $d$-dimensional space. A RBF model is then a linear combination of basis functions, whose argument is the Euclidean distance between the point $x$ at which the interpolation is made and the $n$ points in the known data set. In other words, the interpolation at an untried site is a sum of contributions from all the known function values, the influence of which is controlled by a basis function that depends on the distance they are from the new site. If $\phi$ is the chosen basis function and $\| \cdot \|$ is used to denote the Euclidean norm, then an interpolation model $s$ has the form

$$s(x) = \sum_{i=1}^{n} \alpha_i \phi(\|x - x_i\|) + p(x)$$  \hspace{1cm} (3)

where $\alpha_i, i = 1, ..., n$ are model coefficients, and $p(x)$ is an optional polynomial. The coefficients are found by requiring exact recovery of the original data, $s_X = f$, for all points in the training data set $X$. Hence the model is an interpolant, and all original solution information is preserved. When the polynomial term is included, the system is completed by the additional requirement

$$\sum_{i=1}^{n} \alpha_i p(x) = 0$$  \hspace{1cm} (4)

which is sometimes referred to as the side condition, for a polynomial that takes the form

$$p = \gamma_0 + \sum_{j=1}^{d} \gamma_j x^j$$  \hspace{1cm} (5)

Setting up a global RBF interpolation then requires a solution to a linear system:

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ f_1 \\ \vdots \\ f_n \end{pmatrix} = \begin{pmatrix} 0 & P \end{pmatrix} \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_d \\ \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$  \hspace{1cm} (6)

where

$$P = \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \\ \vdots & \ddots & \vdots \\ x_1^d & \cdots & x_n^d \end{pmatrix} \quad M = \begin{pmatrix} \phi_{1,1} & \cdots & \phi_{1,n} \\ \vdots & \ddots & \vdots \\ \phi_{n,1} & \cdots & \phi_{n,n} \end{pmatrix} \quad \phi_{i,j} = \phi(\|x_i - x_j\|)$$

Once the coefficients are known, the objective function at an interpolated point is then found by equation\textsuperscript{43}
Many choices are available for the basis function itself, but a radial conditionally positive definite function is chosen to ensure the system has a unique solution. Basis functions are said to be compactly supported when they decay to zero at a given distance from the centre, known as the support radius \( R \). Compactly supported functions have desirable numerical characteristics and modelling behaviour, hence the work presented here concentrates on Wendland’s functions \( \phi_{d,k} \), which are compact functions of minimal degree for a stated number of continuous derivatives \( C^2 \) in \( d \) dimensions. Wendland’s \( C^2 \) (\( \phi_{d,1} \)) in up to three dimensions is given by

\[
\phi_{3,1} = (1 - r)^4 (4r + 1),
\]

where \( r = \| \cdot \| \) denotes the norm and \((\cdot)_+\) refers to the cut-off function

\[
(1 - r)_+ = \begin{cases} 
1 - r, & 1 - r \geq 0, \\
0, & 1 - r < 0.
\end{cases}
\]

The support radius used in the function also acts to control the region of influence of each of the centres. A larger support radius allows each sample point to influence the interpolation at a given point from a greater distance away, and in general leads to a smoother interpolation. For a compact function, it defines a hypersphere around each sample point outside of which the value of \( \phi \) is zero, and inside of which the point has some influence. The support radius scales the basis function argument as follows:

\[
r_{\text{scaled}} = \frac{r}{R},
\]

hence the distance between two points is then the scaled Euclidean distance

\[
\| x - x_i \|_{\text{scaled}} = \frac{1}{R} \sqrt{\sum_{j=1}^{d} (x_j - x_{i,j})^2}.
\]

V.B. VG Surrogate Model

The surrogate model is constructed using Wendland’s C2 function with a support radius sufficient to cover the entire design space. Selected projections for the objective function \( J \) are presented in figure 6 for the design space \([0, 1]^5\).

VI. VG Design Optimization

The use of a surrogate model allows very quick evaluation of an interpolation of an otherwise expensive, high-fidelity design space. The high fidelity nature of the data used to produce the interpolation is, however, preserved by building a surrogate model using RBFs. The next stage of this work is the optimization of the VG design based on the five parameters outlined as being important to VG design.

One option for the VG design optimization is to use the simulation approach developed in this work within an optimization framework, which is known as full-fidelity aerodynamic shape optimization. This approach, however, is expensive as very large numbers of simulation evaluations are required during the process. A less expensive approach, and one that exploits the use of the surrogate model, is surrogate-based optimization and is the approach used in this work. This optimization approach substitutes a full simulation evaluation to obtain an objective function value for a given set of inputs with a surrogate model evaluation to obtain an interpolated approximation to the true value of the objective function for the given inputs. The surrogate evaluation is simply given by equation 3, and represents a very cheap way of obtaining responses to the output for the given inputs.

To interrogate the surrogate design space, an optimization algorithm is required. One choice of optimization algorithm is a gradient-based method, however, these are prone to termination in a local minimum that is not necessarily the global minimum solution. Gradient-based approaches are well suited to optimization problems with expensive objective function evaluations, such as full-fidelity aerodynamic shape optimization, for example. When multimodality (multiple local minima) is an issue, a global optimization algorithm is often more effective at locating the globally optimal solution. These algorithms tend to be agent-based, such as particle swarm or gravitational search, or can also be evolutionary-based, such as differential
Figure 6: Surrogate model projections (blue represents lower objective, $J$)
evolution\textsuperscript{22} These global methods usually require many more objective function evaluations compared to gradient-based methods and are therefore suited to problems that have cheap objective function evaluations.

An objective function evaluation in this work is represented by a surrogate evaluation, which is very cheap to evaluate. Furthermore, due to the unknown nature of the modality in the design space, a global optimization algorithm has been developed for optimizing the VG surrogate design space. The global algorithm that has been developed is based on the differential evolution algorithm presented by Storn and Price\textsuperscript{23} and is outlined in section VI.A.

In addition to searching the surrogate design space for an optimum solution directly, a complementary approach interrogates the surrogate model in order to both refine regions of interest and ensure the accuracy of the surrogate design space. Regions of interest in this instance being local minima within the design space. An efficient method of improving surrogate-based accuracy and refinement is adaptive sampling. Adaptive sampling relies on reconstructing/updating the surrogate model across a number of iterative cycles by updating the surrogate model with newly determined adaptive sample points. A kriging based adaptive strategy was compared to a similar RBF adaptive strategy in Mackman \textit{et al}.\textsuperscript{54} These utilise the primarily space filling MaxMSE (largest value of error predicted by the mean squared error function, otherwise referred to as entropy\textsuperscript{64} and a statistical optimization function (expected improvement function). Expected improvement balances local and global search criteria\textsuperscript{63} An alternative RBF adaptive approach that incorporates both local gradient-based refinement and space-filling properties has also been developed\textsuperscript{12} This approach is presented in section VI.B.

VI.A. Differential Evolution

Differential evolution (DE) is a swarm intelligence algorithm built around the concept of evolutionary mechanics. Like the other swarm intelligence algorithms DE takes a number of agents (parameter vectors) which are often termed chromosomes, and follows three steps to advance the optimization algorithm:

1. \textit{Mutation}: generate a mutated candidate solution by small, random changes
2. \textit{Crossover}: ‘breeding’ of mutated and parent solutions to produce a child
3. \textit{Selection}: accept the new child based on its fitness

The parameter vectors are vectors of length $d$, where each entry represents a position in each of the $d$ design parameters:

$$\mathbf{x}_n = \{x_{n1}, x_{n2}, \ldots, x_{nd}\}^T \quad (11)$$

The mutation stage involves the production of a new, candidate solution to introduce variability and exploration into the algorithm. The candidate/mutated solution for the $n$-th parent/target solution is called the donor vector, $\mathbf{v}_n$, and is produced by combining three existing parameter vectors:

$$\mathbf{v}_n = \mathbf{x}_{n1} + F(\mathbf{x}_{n2} - \mathbf{x}_{n3}) \quad (12)$$

where $r_1$, $r_2$ and $r_3$ ($r_1 \neq r_2 \neq r_3 \neq n$) are random integers chosen from the range $[1, N]$, for a population of $N$ parameter vectors (or agents), for each parent vector, and $F$ is a scalar that is typically in the interval $[0, 1]$ and is chosen to either maximize or minimize exploration away from the existing agents.

The crossover stage is used to enhance diversity in the population by combining aspects of the given parent and donor vectors. For the $d$-th entry in the parent ($x_{n1}^d$) and donor vectors ($u_{n1}^d$), the $d$-th entry in a trial/child vector ($u_{n2}^d$) is produced where each entry is from the donor vector based on a probability $CR$, otherwise it is from the parent vector. The ensure there is guaranteed variability and that $u_n$ gets at least one entry from $\mathbf{v}_n$, a random entry from $\mathbf{v}_n$ in always introduced into $u_n$.

The final stage of the algorithm for each of the parameter vectors is to assess whether the new child is of sufficient fitness to replace the given parameter vector. If the fitness of the child vector is better than its parent then the child replaces the parent.

Multiple further additions have been presented for DE, and in the review by Neri and Tirronen\textsuperscript{24} many of these were tested for cost and performance comparison. A small addition that is simple to make to DE to considerably improve its performance is population size reduction (PSR)\textsuperscript{65} which splits the overall number of iterations of the optimizer into a number of stages where the number of parameter vectors are reduced by
a half at each stage. Only the parameter vectors with the best fitness are kept after each stage which has the effect of managing the loading between exploration with many vectors at the beginning of the optimization, and exploitation towards the end of the optimization.

The DE algorithm was run with constants commonly found in the literature; 50 particles, crossover probability 0.3, mutation rate 0.7, 300,000 function evaluations and 4 stages for the PSR.

VI.B. Adaptive Sampling

An adaptive sampling method requires an initial set of samples to first build an interpolation, and then updated sampling points are added based on a compromise between space-filling updates and local refinement in non-linear regions. This non-linearity is determined by the Laplacian. A smooth separation function quantifies the sample spacing in order to assess space-filling updates. A criterion, $C$, was proposed in [12] which merges these two criteria:

$$ C = (|\nabla^2 J| + \epsilon)(1 - h)^2 \quad 0 \leq h \leq 1 $$

where $\epsilon$ is a small number, sufficient to avoid $C = 0$ when $|\nabla^2 J| = 0$. The variable $h$ is the space-filling function, such that $1 - h$ grows away from previously sampled locations. However, this criterion $C$ was proposed in order to improve information around both maxima and minima within the surrogate model (hence the magnitude of the Laplacian). If only a minimum is desired (positive Laplacian) or maximum (negative Laplacian) then the $C$ function is adjusted appropriately. Larger values of $C$ indicate new potential sample locations. Either a grid evaluation can be conducted to determine maximum $C$, and hence the next sample location, or a more robust and efficient strategy utilising the global optimization algorithm outlined in the previous subsection is undertaken.

The space-filling (or separation) function, $h$, is defined as a smooth function, for blending with the local refinement component of $C$. In this function, $h = 1$ at a sample point, and diminishes as distance increases away from the sample point. This is produced using an RBF kernel with a compactly supported basis function and small support radius. The support radius used in this work was determined as a function of two distances, the minimum distance ($d_1$) and supremum minimum distance ($d_2$) between sample points.

$$ Sr_h = R_h \sqrt{(d_1 d_2)} $$

where $R_h$ is a scaling factor, which in this work was $R_h = 1$. However, it was found that the adaptive sampling approach outlined above was insufficient to progressively refine the solution without significant adaptive sample points. Therefore the adaptive sampling approach was modified with the addition of an exploitative term

$$ C_b = (|\nabla^2 J| + \epsilon)(1 - h)^2 b^3 $$

where $b$ is a distance based function and $0 \leq b \leq 1$. The $b$ function is also smooth and produced via an RBF kernel in the same manner as $h$. Instead of using all sampled points, the function merely utilises a subset of $p$ best fitness values taken from the full known sample set, where $p$ is the number of samples in the subset. The support radius for the $b$ function is the maximum distance between a subset sampled point to the full sample set. The $b$ function will be have a value of $b = 1$ at any subset sample point and will decay away from the subset of $p$ points to zero at the furthest reaches of the domain (relative to the subset points). Thus the original adaptive sampling approach can be seen as primarily explorative (when $b = 1$ throughout the design space), whilst the latter is exploitative ($0 \leq b \leq 1$). This can be demonstrated using two analytical functions in the following subsections.

VI.B.1. Franke’s Function

Franke’s function is a bivariate function, consisting of two Gaussian peaks of differing heights and a smaller trough. In this instance, the adaptive sampling is undertaken using a grid evaluation seeking only the maxima location, with the trough refinement ignored. The function is defined using
\[ f(x) = 0.75 \exp \left( -\frac{(9x_1 - 2)^2}{4} - \frac{(9x_2 - 2)^2}{4} \right) + 0.75 \exp \left( -\frac{(9x_1 + 1)^2}{49} - \frac{9x_2 + 1}{10} \right) \]
\[ + 0.5 \exp \left( -\frac{(9x_1 - 7)^2}{4} - \frac{(9x_2 - 3)^2}{4} \right) - 0.2 \exp \left( -(9x_1 - 4)^2 - (9x_2 - 7)^2 \right) \]

between \( x_i \in [0, 1] \).

After an initial 9 sample points, 24 adaptive sample points are introduced (sequentially, rather than in batches). The results for using both adaptive sampling techniques is given in figure 7.

![Figure 7](image-url)

(a) \( C \) Approach (2d)  
(b) \( C \) Approach (3d projection)

(c) \( C_\theta \) Approach (2d)  
(d) \( C_\theta \) Approach (3d projection)

Figure 7: Comparison of \( C \) and \( C_\theta \) adaptive sampling approaches using Franke’s Function

From the figure, it is clear that the \( C_\theta \) approach is highly concentrated around the primary peak. Less obviously, it is also more rapid in identifying the exact peak location during the adaptive cycle.

VI.B.2. *Rastrigin Function*

The two-dimensional variant of the \( n \)-dimensional Rastrigin function is a more complex problem, with a plethora of local minima and maxima. It is defined by

\[ f(x) = -An - \sum_{i=1}^{n} \left( x_i^2 - A \cos (2\pi x_i) \right) \]
where \( A = 10, n = 2, x_i \in [-5.12, 5.12] \). The global maximum \( f(x) = 0 \) is located at \( x = 0 \). After an initial 16 sample points, 24 adaptive sample points are introduced sequentially. As shown in figure 8, both adaptive approaches capture the overall global maximum, although the \( C_b \) function has again clearly covered all potential local maxima within proximity.

![Graphs showing adaptive sampling approaches using Rastrigin's Function](image)

Figure 8: Comparison of \( C \) and \( C_b \) adaptive sampling approaches using Rastrigin’s Function

Both adaptive strategies offer advantages, with \( C \) providing an initial adaptive coverage, and \( C_b \) refining the final solution.

**VI.C. VG Optimization Results**

The results of optimizing the VG surrogate model for the objective function given by equation 1 are given in table 4 along with the clean values and the values that are the lowest objective values from the latin hypercube sample. To increase the adaptive sampling rate, a number of updates per iteration were undertaken in parallel. This only requires updating the separation function before determination of the next update location. The number of updates per iteration and the number of iterations was prescribed in advance. Two rounds of explorative adaptive sampling using five samples in each round were followed by one round of exploitative sampling.

The results clearly demonstrate that the optimized value is a considerable improvement in the separation experienced on the aerofoil and therefore a reduction in buffet may be expected. The optimization process
via differential evolution has also improved the performance of the VG compared to just taking the best sample tested in the latin hypercube, therefore vindicating the optimization process. Comparing to the clean geometry, it is particularly impressive that the process developed here, of design space sampling, surrogate-based optimization and adaptive sampling, has been able to eliminate separation at a highly loaded condition and achieved a decrease in the drag on the aerofoil of 7% relative to best result achieved by the Latin hypercube sampling alone.

Table 4: Design space results

<table>
<thead>
<tr>
<th></th>
<th>J</th>
<th>CD</th>
<th>Separation</th>
<th>h</th>
<th>AR</th>
<th>θ</th>
<th>c</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clean</td>
<td>0.0246</td>
<td>0.0190*</td>
<td>11.1%*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Min. LHC</td>
<td>0.0204</td>
<td>0.0203</td>
<td>0.08%</td>
<td>0.81%</td>
<td>3.84</td>
<td>16.89°</td>
<td>55.6%</td>
<td>8.8</td>
</tr>
<tr>
<td>Optimized (Surrogate)</td>
<td>0.0189</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Optimized (CFD)</td>
<td>0.0189</td>
<td>0.0189</td>
<td>0.0%</td>
<td>0.85%</td>
<td>3.64</td>
<td>8.42°</td>
<td>21.5%</td>
<td>5.8</td>
</tr>
</tbody>
</table>

* Oscillatory result: time-averaged

Streamlines proximal to the VG and trailing edge of the airfoil are given in figure 9. From the figure, the large separation region present in the clean geometry has been entirely eliminated using the optimum VG configuration; see figure 3a for the clean geometry streamlines. Figure 9c shows the vortex extent, as defined by the \( \lambda_2 = 0 \) criterion between the minimum LHC sample and the optimum solution. This criterion denotes the intermediate eigenvalue of the symmetric tensor \( S^2 + \Omega^2 \) where \( \Omega = \frac{1}{2} [\nabla u - (\nabla u)^T] \) is the vorticity tensor and \( S = \frac{1}{2} [\nabla u + (\nabla u)^T] \) the rate of strain tensor. Negative values of the \( \lambda_2 \) criterion denote the vortex. The induced vortex of the optimum design has eliminated separation at the trailing edge of the airfoil, as shown in figure 9b although a small region is still present in the minimum LHC sample.

![figure 9](image)

Figure 9: Induced vortex impact downstream of VG

The high performance of the optimization algorithm can be demonstrated by considering the value of the gradients of the objective with respect to changes in the design variables at the optimized point. An optimum solution for this type of problem is one where the gradient values approach a very small number. Table 5 shows these gradient values and demonstrates the high performance of the optimization algorithm, producing a solution that is optimum to within a very small tolerance, and an improvement on the minimum sample from the Latin Hypercube, which is not at an optimum solution as defined by the gradient.
VII. Conclusions

The quantification of the effects of flow around a tiltrotor-like aerofoil with vortex generators (VGs) and subsequent design of these VGs has been considered. The highly loaded and thick wing of a tiltrotor is particularly susceptible to early onset buffet, so boundary layer control is critical and, hence, upper surface flow control devices are an important consideration. An analysis package has been presented that uses the open source library, OpenFOAM, to obtain high-fidelity viscous flow simulation, and a custom mesh generator to autonomously obtain high quality structured meshes.

Design of counter rotating VGs has been considered using, firstly, a Latin Hypercube sampling approach to obtain design space samples for five variables; VG height, VG aspect ratio, VG angle, chord-wise location, and inter- and intra-pair spacing ratio. A surrogate model that uses radial basis functions has been developed to produce an approximation to a high-fidelity design space but still preserves the important flow physics required for design. An optimization algorithm based on differential evolution, which is a global search algorithm, was used to fully interrogate the surrogate model design space. With the initial sample points defined via an optimised Latin Hypercube, adaptive sampling using a combined objective function Laplacian/space-filling criterion was then employed in order to locate a global optimum VG design for the high loading manoeuvring condition. The objective function blends the requirement for drag minimization with reduction of separation.

The best Latin Hypercube result obtained reduced the separation from 11.1% to 0.08%, but increased the drag by 6.8% relative to the clean geometry. Following the surrogate-based design optimization, the optimum result obtained has eliminated separation entirely at $C_L = 1.3$ whilst decreasing the drag by 7% relative to the minimum Latin Hypercube sample and with a marginal decrease in drag relative to the clean geometry.

Acknowledgements

The authors would like to thank David Tring and Nigel Scrase of AgustaWestland, for their support of and contribution to this work. The financial support, via the HiPerTilt project, of AgustaWestland and Innovate UK (Technology Strategy Board) is also gratefully acknowledged. Finally, the authors would like to acknowledge that this work was carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol.

References


