INTRODUCTION

Given the high cost of CFD and optimization, a prominent area of research today is to find ways to reduce the computational time while retaining the high-fidelity of the analysis. In the area of aerodynamic optimization, the variable fidelity (also called multi-fidelity) method has quickly grown in popularity (Alexandrov et al., 2001; Keane, 2003; Gano et al., 2005; Forrester et al., 2006; Forrester et al., 2007; Nelson et al., 2007; Lehner et al., 2010).

Variable-fidelity and other model management methods have been developed to solve optimization problems that involve simulations with large computational expense (Gano et al., 2005; Huang et al., 2006). In many engineering design problems, differing levels of fidelity can model the system of interest. Higher-fidelity models typically incorporate more detailed physics and are computationally expensive to evaluate than lower fidelity models. The lower-fidelity models are typically much cheaper to evaluate, but designs produced by using these models neglect important physical effects included in the more expensive higher-fidelity models. In aircraft design, Navier-Stokes and Euler equations are examples of two computational models with different fidelity.

Variable-fidelity optimization has emerged as an attractive method of performing, both, high-speed and high-fidelity optimization (Alexandrov et al., 2001; Keane, 2003; Leary et al., 2003; Gano et al., 2005; Forrester et al., 2006; Huang et al., 2006; Forrester et al., 2007; Nelson et al., 2007; Lehner et al., 2010). These algorithms attempt to leverage information from computationally inexpensive low-fidelity models to reduce the time required to converge to the optimum of the high-fidelity function. This is usually accomplished by building a computationally inexpensive surrogate for the high-fidelity model. The surrogate is, often, iteratively optimized and updated as new high-fidelity data become available. In variable-fidelity algorithms, the surrogate for the high-fidelity model usually consists of the low-fidelity model plus a correction term that models the difference between the high- and low-fidelity models, calibrated at selected sample points in the design space (Lehner et al., 2010).

The variable-fidelity method has its roots in the empirical model-building theory – the idea of endowing a surrogate with some discipline related properties to increase its accuracy—and the past three decades have seen rapid increase in its development and use (Simpson et al., 2008). Insightful reviews of the variable-fidelity method can be found in Huang et al. (2006) and Simpson et al. (2008). The method has improved immensely; from its initial form requiring Taylor polynomials (Alexandrov et al., 2001) to its current incarnation using a variety of modern surrogate models like Kriging and neural networks (Leary et al., 2003; Gano et al., 2005); and from using a variety of multiplicative, additive and hybrid corrections (Gano et al., 2005), to the method of Co-Kriging (Forrester et al., 2007). In the optimization context, the method has been used in both derivative-based (Alexandrov et al., 2001) and derivative free (Booker et al., 1999) optimization.

The authors have previously demonstrated a Variable-Fidelity Optimization (VFO) framework for...
perform parallel computations. This study presents a based methods do making it less prone to being trapped are also compared to the results of direct optimization, the HF solver, alone, was coupled to the results of the optimization. The VF optimization results next, the VFO framework is presented followed by the technology pieces used in the design space (Zahir and Gao, 2011). The authors’ framework worked well: The lift-to-drag ratio, $K$, was within 6% of the HF optimum and provided significant savings in computation time. Detailed results can be found in Zahir and Gao (2011). In the quest for greater efficiency and accuracy, the framework is modified to perform parallel computations. This study presents a minimum drag VFO of a transonic airfoil, at constant lift, subject to thickness and pitching moment constraints using several low fidelity solvers.

This study is arranged as follows: First the major technology pieces used in the design are described; next, the VFO framework is presented followed by the results of the optimization. The VF optimization results are also compared to the results of direct optimization, where the HF solver, alone, was coupled to the optimizer to find the optimum result.

MATERIALS AND METHODS

Flow solver: In this study, the RAE-2822 airfoil is chosen for optimization. An indigenously developed 2D compressible Navier-Stokes solver-using the LU-SGS time stepping scheme, the Roe upwind scheme and multigrid acceleration-capable of being used in, both, Euler and Reynolds-Averaged-Navier-Stokes (RANS) mode (Su et al., 2008) is used as the flow solver. Euler and low resolution (small grid) RANS solvers are used as the LF solvers, while RANS with the K-ω turbulence model is used as the HF solver.

The Euler and RANS solvers use a C-type mesh extending 20 chord lengths downstream of the trailing edge. The first grid line is $2 \times 10^4$ units from the airfoil surface for the RANS solver and 0.01 units for the Euler solver. The HF solver uses a grid size of 216 cells around the airfoil and 44 cells normal to the airfoil (216x44 grid), selected after obtaining good agreement in the surface pressure distribution and aerodynamic coefficients at transonic conditions (Mach number of 0.729, Reynolds number of $6.5 \times 10^6$ and an angle of attack of $2.31^\circ$) as reported in Cook et al. (1979) and Slater (2002).

Optimization algorithm: A Genetic Algorithm (GA) is used to perform the optimization in this study. GA searches from multiple points in the design space, instead of moving from a single point like gradient-based methods do making it less prone to being trapped by local optima. This makes it particularly suitable for aerodynamic optimization where the function landscape is often multimodal and nonlinear because the flow field is governed by a system of nonlinear partial differential equations (Oyama et al., 2000). Furthermore, GA works on function evaluations alone and does not require derivatives or gradients of the objective function. These features make it a robust global optimization algorithm.

The GA implementation in MATLAB and its optimization toolbox is used to perform the optimization.

Sample plan: As with all surrogate-based methods, to approximate a function $f$ we start with a set of sample data-computed at a set of points in the domain of interest determined by a sampling plan. Selection of the sample points is a very important step towards creating a good surrogate model. If the sample plan model is too sparse or does not contain the interesting features of the design space, the resulting model will fail to resolve desirable global features. In order to improve the surrogate it is necessary to incrementally add points in an intelligent way such that the generated surface converges toward the true surface.

When dealing with large, complex design spaces, it is often unclear how many points may be necessary to resolve key features with a response surface. To get the maximum amount of information out of a minimum number of points with no a priori knowledge of the design space requires uniform sampling.

Several sampling methods are capable of producing relatively uniform samples, e.g. Latin Hypercube sampling (Forrester et al., 2008), however not many methods allow incremental uniform sampling. Latin Hypercube sampling requires a priori knowledge of how many points are desired in order to divide the domain into the appropriate number of hypercubes. This method creates nearly uniform point distributions, but requires a completely new set of data if additional points are desired. Another sampling method known as the Sobol Sequence (Sobol, 1967) has good space filling properties and allows incremental uniform sampling (Nelson et al., 2007; Forrester et al., 2008). This method is, thus, adopted in this study to create the sample plan. Since the number of sample points required to obtain an accurate surrogate is generally unknown a priori, an initial sample of $10n_{var}$ (where $n_{var}$ is the number of design variables) is created following the suggestion of Jones et al. (1998).

SURROGATE MODEL

The surrogate model used in the study is Kriging-an approximation technique that has received much
attention in recent years (Sacks et al., 1989; Jones et al., 1998; Keane, 2003; Leary et al., 2003; Gano et al., 2005; Forrester et al., 2006; Huang et al., 2006; Forrester et al., 2007; Nelson et al., 2007) named after the pioneering work of the South African mining engineer D.G. Krige and introduced in engineering design work after the seminal paper by Sacks et al. (1989).

For brevity, the Kriging equations are not mentioned here. Forrester et al. (2008) contains more detail about Kriging.

The Kriging model in this study uses a constant regression term and a Gaussian correlation model. The p term is 2 for all dimensions and θ is optimized in the range $10^2 \leq \theta \leq 2000; i = 1, 2, \ldots, n_{var}$. The surrogate model is created using the surrogates toolbox (Viana, 2010) for MATLAB.

**Design variables:** The design process begins with an initial airfoil. The airfoil geometry is then modified by adding smooth perturbations in the form of the Hicks-Henne bump functions (Hicks and Henne, 1978). Defining normalized airfoil coordinates $\psi = x/c$, $\zeta = z/c$ and c as the chord length, the Hicks-Henne shape function modifies airfoil geometry by adding a linear combination of shape functions, $F_i$ and weighting coefficients, $x_i$ as follows:

$$
\zeta = \zeta_{x,value} + \sum_{i=1}^{n_{fvars}} x_i F_i(\psi),
$$

$$
F_i(\psi) = \left[ \sin \left( \pi \frac{\psi - \log \frac{\psi - \log_{10} a}{\log_{10} b}}{\log_{10} a - \log_{10} b} \right) \right]^{t_i}, 0 \leq \psi \leq 1
$$

Here $t_i$ locates the maximum point of the bump and $t_j$ controls the width of the bump. The design variables are the coefficients $x_i$ multiplying the various Hicks-Henne bump functions.

In this study, 7 bump functions are used for the upper and lower surface of the airfoil, resulting in a total of 14 design variables. The points $t_i$ are linearly spaced between 0 and 0.94. The range is terminated a little before the trailing edge to prevent the upper and lower edges from crossing each other and creating unrealistic geometries. A value of 10 is used for $t_j$ following Castonguay and Nadarajah (2007) recommendation.

To prevent large changes to the geometry, upper and lower bounds are set on $x_i$. These are:

$$
-0.005 \leq x_i \leq 0.005, i = 1, 2 \text{ and } 13,14
$$

$$
-0.01 \leq x_i \leq 0.01 \text{ otherwise}
$$

**Fitness function and constraints:** This is a single objective optimization problem. The airfoil is optimized for a transonic Mach number of 0.729 and a fully turbulent Reynolds number of $6.5 \times 10^6$. The objective is to minimize the drag, $c_d$ at a constant lift, $c_l=0.6\pm0.01$, subject to thickness and pitching moment constraints. RAE2822 is used as the initial airfoil to start the optimization. The optimization problem is stated as follows:

Minimize:

$$
c_d(x)
$$

subject to:

$$
x_i \leq x \leq x_i,
0.59 \leq c_l \leq 0.61
$$

$$
c_m \geq -0.122
$$

$$
\left(\frac{f(x)}{c_l}\right)_{max} \geq 12.11\%
$$

Constant lift is maintained by varying the angle of attack, $\alpha$, of the airfoil. At low to moderate values of $\alpha$, i.e., before flow separation and stall, $c_l$ varies linearly with $\alpha$. During each evaluation of the fitness function, the airfoil is first simulated at an initially guessed $\alpha_1$ and $\alpha_2$. A third simulation at $\alpha_3$, found from a linear interpolation through $(\alpha_1, c_1)$ and $(\alpha_2, c_2)$, is then used to attain the desired lift. This procedure effectively removes the $c_l$ constraint from the optimization by making it a condition which each simulation must satisfy. The optimization is simplified both by the removal of the constraint and by reducing the dimensionality of the problem, since $\alpha$ is not used as an optimization variable. Other constraints are imposed by adding penalty terms to the fitness function:

$$
f(x) = c_d |c_{lift}| + \max[0, (t/c) - (t/c)_{max}] + \max[0, c_m - c_m_{min}] - 0.122
$$

where $(t/c)_{min}$ is the minimum allowable thickness of $12.11\%$ (the $(t/c)_{max}$ of the initial airfoil) and $c_m_{min}$ is -0.122

**The VFO framework:** A variable fidelity prediction predicts the function response in 2 steps:

- The low fidelity function is evaluated to obtain an estimate
- The estimate is corrected by performing a high fidelity simulation to obtain a better estimate to the function and apply a correction to the low fidelity prediction. The correction between the high and low fidelity functions is modeled by a surrogate model developed by sampling the function at a few points. To evaluate the goodness of fit over the entire function landscape, the RMSE and $R^2$ of the surrogate is calculated for a validation dataset generated using a uniform Latin Hypercube sample consisting of $n_{test} = n_{var} \times 10$ points.
Figure 1 shows the flowchart of the VFO framework used in this study. If the initial number of points used to construct the surrogate do not result in an accurate surrogate, the surrogate is updated by adding more points until some accuracy criteria is satisfied. The surrogate is then directly coupled to the GA to determine the optimum.

It has been reported that when a surrogate is used for fitness evaluation, it is very likely that the evolutionary algorithm will converge to a false optimum (Jin, 2005). A false optimum is an optimum of the approximate model, which is not one of the original fitness function. To avoid this, the VF optimum point is evaluated by the HF solver every 10 generations. If the VF lift coefficient, $c_l$, agrees with the HF $c_l$ prediction within a 0.01 tolerance and the drag coefficient within 0.0001, the optimization is continued. Otherwise, the surrogate model is updated by performing HF evaluations on the GA population and refitting the surrogate before continuing the optimization.
Optimization is terminated once the cumulative change in the fitness function value over 50 generations is less than or equal to 1.0\textsuperscript{6}, or when 100 GA generations are completed (whichever comes sooner).

**Parallelization:** MATLAB’s distributed computing toolbox provides a scheduler to distribute multiple tasks on several cores of the same computer, or on different computers. Although the CFD code used is a serial code, the unique individuals of each population are evaluated in parallel on several computers (and cores) achieving significant time savings. A heterogeneous cluster of 7 dual-core IBM-PC computers clocked at 2.5 GHz and 3.0 GHz (resulting in 14 nodes) and a head node are used in this study.

**RESULTS AND DISCUSSION**

Use of Euler and RANS solvers with several grid sizes was investigated previously by Zahir and Gao (2011) for use as the low-fidelity solvers. The surface pressure distribution on the RAE-2822 at Mach number of 0.729, Reynolds number of 6.5x10\textsuperscript{6} and an angle of attack of 2.31° are shown in Fig. 2 (Zahir and Gao, 2011). Results of the HF solver (NS 216x44) are also shown. All Navier Stokes solvers use the K-ω turbulence model and an initial grid line spacing of 2x10\textsuperscript{6} units from the airfoil surface. The Euler solvers use an initial grid line spacing of 0.01 units from the airfoil surface. Difference between the aerodynamic coefficients of the low-fidelity solvers and the HF solver is shown in Fig. 3. The computation time for one run is also shown (Zahir and Gao, 2011).

Three LF solvers are selected for the VF surrogate in this study to cover the entire spectrum from slow to fast and less accurate to more accurate: Euler 60x20, Euler 160x24, RANS 160x24. Figure 4 shows the
Table 1: Aerodynamic coefficients of the VF optimized airfoils using several LF Solvers. High fidelity evaluations at the VF optimum points are also shown along with total optimization time. Results of direct optimization using the HF solver alone are also shown for reference.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>VF optimum Result</th>
<th>HF evaluation at VF optimum point</th>
<th>Surrogate creation Time (Hours)</th>
<th>Opt. time (Hours)</th>
<th>Total time (Hours)</th>
<th>Hours saved (compared to direct opt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VF, Euler 60x20, 140 samples</td>
<td>0.598 0.00628</td>
<td>0.599 0.00630 -0.095 95.08</td>
<td>0.57</td>
<td>0.91</td>
<td>1.77</td>
<td></td>
</tr>
<tr>
<td>VF, Euler 160x24, 140 samples</td>
<td>0.598 0.00517</td>
<td>0.598 0.00650 -0.089 92.01</td>
<td>12.16% 0.41</td>
<td>0.81</td>
<td>1.22</td>
<td>1.46</td>
</tr>
<tr>
<td>VF, NS 160x24, 140 samples</td>
<td>0.599 0.00610</td>
<td>0.598 0.00609 -0.096 98.21</td>
<td>12.12% 0.50</td>
<td>1.02</td>
<td>1.52</td>
<td>1.16</td>
</tr>
<tr>
<td>HF, NS 216x44 (direct optimization)</td>
<td>0.598 0.00603</td>
<td>0.598 0.00603 -0.096 99.18</td>
<td>12.12% 0.41</td>
<td>1.52</td>
<td>2.68</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Comparison of results obtained from serial and parallel runs, benefits of parallel computation

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Aerodynamic coefficients</th>
<th>Optimization time (hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serial Run (Zahir and Gao, 2011)</td>
<td>0.600 0.00643 -0.123 93.45</td>
<td></td>
</tr>
<tr>
<td>Parallel Run (Zahir and Gao, 2012)</td>
<td>0.610 0.00656 -0.122 93.03</td>
<td>12.13% 2.31</td>
</tr>
</tbody>
</table>

Fig. 5: Surface pressure distributions on the optimum airfoils. Results were calculated using the HF solver on the geometry produced by the VF optimization algorithm. The pressure distribution obtained by direct optimization using the HF solver is also shown.

RMSE and R² values for the Kriging surrogate fitted to the difference between the HF and LF response, ∆f and compared to the validation dataset.

The surrogates yield high R² values and are used to perform the VF optimization. The number of sample points used to create the VF surrogate are such that R²>0.9. In this case, it happens to be the minimum number of training points, 140. The optimization results are given in Table1 with optimization time. HF evaluations of the aerodynamic coefficients at the VF optimums are also reported.

It is clear from Table 1 that the RANS 160x24 solver yields the best lift-to-drag ratio, k, within 1% of the direct optimization result. Other solvers perform well too; yielding k within 7% of the direct optimization result, while still providing a significant saving in computation time. All solvers also satisfy the thickness constraint (t/c) max ≥ 12.11%.

The benefits of parallel computation are shown in Table 2 for a similar optimization problem, but without the pitching moment constraint, previously reported in Zahir and Gao (2011) and Zahir and Gao (2012). The parallel computations are 24 times faster than the serial computations a very significant improvement.

The surface pressure distributions on the optimum airfoils are shown in Fig. 5 and the airfoil geometries are shown in Fig. 6. It is seen that the optimum pressure
distributions and airfoil geometries produced using different LF solvers are quite similar. This is also reflected in the aerodynamic coefficients shown previously in Table 1.

CONCLUDING REMARKS

The VFO method used in this study is, both, efficient and capable of finding the optimum that closely agrees with the results of high-fidelity optimization alone. Modification of the authors’ original VF framework to incorporate parallel evaluation yields significant time savings. Euler and Navier-Stokes solvers evaluated on low resolutions grids are good candidates for LF solvers and produce results close to the HF optimum with significant time savings.

REFERENCES


