Efficient Global Optimization of Helicopter Rotor Blades for Vibration Reduction in Forward Flight

Bryan Glaz*, Peretz P. Friedmann† and Li Liu‡

Department of Aerospace Engineering, The University of Michigan, Ann Arbor, MI, 48109, USA

The effectiveness of surrogate based optimization of helicopter rotor blades for vibration reduction using kriging global approximations is investigated. The search for the optimal design is conducted using two methods: (1) the “one-shot” approach in which the surrogate is optimized directly and (2) a method based on the expected improvement function which seeks to find optimal designs while reducing the uncertainty in the surrogate. The results indicate that the more global search based on the expected improvement function can lead to more optimal rotor blade designs with less function evaluations than the “one-shot” approach, and thus is more efficient. In addition, the results indicate that the application of parallel computation to the expected improvement function based method aids in finding more optimal blade designs even if the kriging approximations are not accurate everywhere in the design space.

Nomenclature

- $c$: Blade chord
- $C_W$: Helicopter weight coefficient
- $C_{d0}$: Blade profile drag coefficient
- $C_{df}$: Flat plate drag coefficient
- $D$: Vector of design variables
- $EIF(x)$: Expected improvement function
- $f(x)$: Assumed polynomials which account for the ‘global’ behavior in kriging
- $F_{4X}, F_{4Y}, F_{4Z}$: 4/rev hub shears, non-dimensionalized by $m_0\Omega^2R^2$
- $g(D)$: Constraints
- $J$: Objective function
- $J_P$: Mass polar moment of inertia of the rotor
- $m_0$: Mass per unit length at the blade root
- $m_{ns}$: Non-structural mass located at the shear center
- $M_{4X}, M_{4Y}, M_{4Z}$: 4/rev hub moments, non-dimensionalized by $m_0\Omega^2R^3$
- $N_b$: Number of rotor blades
- $N_c$: Number of behavior constraints
- $N_{dv}$: Number of design variables
- $N_{sp}$: Number of sample points
- $p_k, \vartheta_k$: Fitting parameters in kriging corresponding to the $k^{th}$ design variable
- $R$: Blade radius
- $R_{krg}$: Spatial correlation matrix used in kriging
- $R_{krg}(\cdot)$: Spatial correlation function in kriging
- $r_{krg}(x)$: Spatial correlation vector in kriging
- $s^2(x)$: Mean square error in kriging surrogate

*Ph.D. Candidate, Student Member AIAA.
†François-Xavier Bagnoud Professor, Fellow AIAA, AHS
‡Postdoctoral Researcher, Member AIAA.
I. Introduction

Vibration is one of the most critical concerns in the design of modern rotorcraft. Stricter demands for enhanced performance, comfort, and customer acceptance require designs with reduced vibration levels. In helicopters, the dominant source of vibrations is the rotor, which transfers vibrations to the rotor hub and fuselage primarily at a harmonic of $N_b/\text{rev}$, where $N_b$ is the number of blades.

During the last 25 years, two principal approaches to vibration reduction have emerged. The first approach is passive and uses structural/multidisciplinary optimization for reducing vibrations,\textsuperscript{1-4} while the second approach utilizes active control methods.\textsuperscript{5,6} This paper focuses on the passive approach. In the passive approach, the vibration reduction problem is formulated as a mathematical optimization problem subject to appropriate constraints. The objective function consists of a suitable combination of the $N_b/\text{rev}$ hub shears and moments that are computed from an aeroelastic response code; the constraints are blade stability margins, frequency constraints, an autorotation constraint, and constraints associated with the blade geometry. The design variables can be dimensions of the blade cross-section, mass and stiffness distributions along the span, or geometrical parameters which define advance geometry tips. Typical levels of vibration reduction achieved with passive approaches have been in the range of 50-60% relative to a baseline design.

Due to the complex aerodynamic environment in the rotary-wing problem, aeroelastic response simulations needed for vibratory load calculations are computationally expensive and therefore numerous evaluations of the vibration objective function are costly. Therefore, direct combination of the objective function generated by the aeroelastic response simulation with traditional optimization algorithms requires excessive computational effort. Moreover, traditional optimization search algorithms can converge to local optima, which are known to occur in this class of problems.

To overcome these obstacles, approximation concepts have been used. A widely used approach in the helicopter literature is to approximate the vibration objective function and constraints using Taylor series expansions about local design points.\textsuperscript{7} The derivatives needed for the Taylor series are calculated using
difference formulas, or analytical sensitivity derivatives. These approximations to the objective function and constraints are used to replace the actual problem with an approximate one that is used in conjunction with the optimizer to obtain an optimal design with a reduced number of expensive function evaluations. Examples of this method can be found in Refs. 2 and 8, where vibration reduction of composite rotor blades with advanced geometry tips in forward flight was studied. The disadvantages of this method are that it is a local approximation in the vicinity of a design point and a local search of the design space is conducted. Even when augmenting this method with move limits or a trust region strategy, convergence is only guaranteed to a local optimum.

An alternative to the local Taylor expansion method is to use global approximation methods; i.e. methods which try to capture the behavior of a function over the entire design space. In one such vibration reduction study, Ganguli used a 2nd order polynomial global approximation of the vibration objective function and achieved 30% vibration reduction. The 2nd order polynomial was found to be accurate only in the vicinity of the baseline design.

In addition to polynomials, there is a class of global approximation methods based on a stochastic process, which yield probabilistic measures of the uncertainty in the approximation (or surrogate) that can be utilized in a global search of the design space. Booker et al. used such an approximation method, known as kriging. In Ref. 15, surrogate methods were applied to minimization of helicopter vibration, using 31 design variables to characterize the rotor blade. The cross-sectional design variables were mass, center of gravity offset from the elastic axis, and the blade stiffnesses. It is important to note that the principal focus of Ref. 15 was the effectiveness of the kriging surrogates, and therefore accurate modeling of the aerodynamic environment of a rotor blade during flight was not considered to be important. Consequently, accurate free wake models were sacrificed for a computationally less expensive prescribed wake model. Furthermore, no constraints were placed on the aeroelastic stability of the blade. Thus, the model of the helicopter vibratory loads was not sufficiently reliable to produce a realistic blade design.

In the study by Glaz et al., various global approximation methods were employed to achieve vibration reduction. The study was based on a comprehensive helicopter simulation code which utilizes advanced modeling techniques such as free wake modeling. Eight design variables were used to characterize the blade’s spanwise mass and stiffness distributions, including blade cross-sectional dimensions and non-structural masses. Surrogates were created to replace the computationally expensive vibration objective function but not the constraints since they were relatively inexpensive to calculate. The results showed that kriging was the most accurate global approximation method considered and led to the best design, which produced 67% vibration reduction. The surrogates were optimized directly even though the approximations were not accurate everywhere in the design space. Although significant vibration reduction was achieved, the focus of the study was on the effectiveness of the approximation methods to capture the behavior of the vibratory loads, and not on the search of the design space for the global optimum.

It is possible to further reduce the number of function evaluations needed to find the optimum by employing a more “intelligent” search of the design space which selects additional search points where (a) there is a high probability of improvement over the current best design and/or (b) there is much uncertainty in the surrogate’s prediction. Such a method, known as Efficient Global Optimization (EGO), is discussed in Refs. 14, 17, 18. The advantages of EGO, which utilize the expected improvement of stochastic process based approximations, are threefold: fewer “true” function evaluations (thus fewer expensive aeroelastic simulations), the formulation is conducive to parallel computing, and facilitation of a more “global” search of the design space.

The overall objective of this paper is to show that the number of function evaluations needed to find an optimal design can be reduced further by combining expected improvement based search methods with kriging approximations, as opposed to optimizing the surrogate directly (as in Ref. 16). To achieve this objective, optimization of helicopter rotor blades for vibration reduction in forward flight will be considered. The helicopter vibration reduction problem represents an extremely complex and computationally expensive application of the EGO method.

This paper is organized as follows: Section II gives an overview of the helicopter simulation used to calculate vibrations, followed by the formulation of the optimization problem in Section III. The next section describes the method for creating global approximations of the vibration objective function using kriging interpolation and the procedure for using the expected improvement function to search the design space for the optimal design. Section V presents results obtained by applying the optimization method to a helicopter rotor blade at a certain flight condition, and Section VI summarizes the conclusions of this work.
II. Overview of the Aeroelastic Response and Stability Analysis

The simulation code used in this study is based on a comprehensive aeroelastic analysis code. The aeroelastic response analysis can represent the behavior of hingeless rotor blades as shown in Fig. 1, with actively controlled flaps; as well as blades with advanced geometry tips shown in Fig. 2. The key ingredients of the aeroelastic response analysis are: (1) the structural dynamic model, (2) the unsteady aerodynamic model and (3) a coupled trim/aeroelastic response procedure that is required for the computation of the steady state blade response. The aeroelastic response analysis and an overview of the aeroelastic stability in hover calculation are described next.

![Coupled Flap-Lag-Torsion Dynamics](image1)

**Figure 1.** Helicopter rotor blade with trailing edge flaps.

![A blade with advanced geometry tip](image2)

**Figure 2.** A blade with advanced geometry tip.

A. Structural Dynamic Model

The structural dynamic model is based on an analysis developed by Yuan and Friedmann which is capable of modeling composite blades with transverse shear deformations, cross-sectional warping, and swept tips. This study is limited to the behavior of isotropic blades with spanwise varying properties. The equations of motion are formulated using a finite element discretization of Hamilton’s principle, with the assumption that the blade undergoes moderate deflections. The beam type finite elements used for the discretization have 23 nodal degrees of freedom. Normal modes are used to reduce the number of structural degrees of freedom. In this study, eight modes are used: the first 3 flap modes, first 2 lead-lag modes, first 2 torsional modes, and the first axial mode.

B. Aerodynamic Model

The attached flow blade section aerodynamics are calculated using a rational function approach (RFA). The RFA approach is a two-dimensional unsteady time-domain theory that accounts for compressibility as well as variations in the oncoming flow velocity. This two-dimensional aerodynamic model is linked to an enhanced free-wake model which provides a non-uniform inflow distribution at closely spaced azimuthal steps. Although the simulation code can also account for dynamic stall at high advance ratios, dynamic stall was not considered in this paper because the vibration levels being approximated are those due to blade vortex interaction (BVI), which occurs at low advance ratios.

C. Coupled Trim/Aeroelastic Response

The combined structural and aerodynamic equations form a system of coupled ordinary differential equations that are cast into first order state variable form and integrated in the time domain using the Adams-Bashforth predictor-corrector algorithm. A propulsive trim procedure, where six equilibrium equations (three forces and three moments) are enforced, is used in this study. The trim equations are solved in a coupled manner with the aeroelastic equations of motion. The vibratory hub shears and moments are found by integrating the distributed inertial and aerodynamic loads over the entire blade span in the rotating frame, then transforming these loads to the hub-fixed non-rotating system, and summing the contributions.
from each blade. In the process, cancellation of various terms occurs and the primary components of the hub shears and moments have frequency of $N_b/\text{rev}$, which is known as the blade passage frequency.

D. Aeroelastic Stability in Hover

The process for determining the hover stability of the blade is similar to that used by Yuan and Friedmann, and is summarized below:

1. The non-linear static equilibrium solution of the blade is found for a given pitch setting and uniform inflow, by solving a set of nonlinear algebraic equations. Note that uniform inflow is used only in the hover stability calculation. The forward flight analysis employs a free-wake model for inflow calculation.

2. The governing system of ordinary differential equations are linearized about the static equilibrium solution by writing perturbation equations and neglecting second-order and higher terms in the perturbed quantities. The linearized equations are rewritten in first-order state variable form.

3. The real parts of the eigenvalues of the first-order state variable matrix, $\lambda_k = \zeta_k + i\omega_k$, determine the stability of the system. If $\zeta_k \leq 0$ for all $k$, the system is stable.

The linearization process in Ref. 8 has been modified to account for the aerodynamic states introduced by the RFA model. Details on the modified linearization process with RFA aerodynamics are provided in Ref. 16.

III. Formulation of the Blade Optimization Problem

The formulation of the blade optimization problem in forward flight consists of several ingredients: the objective function, design variables, and constraints. The mathematical formulation of the optimization is stated as: find the vector of design variables $D$ which minimizes the objective function, i.e. $J(D) \rightarrow \min$, where the objective function consists of a combination of the $N_b/\text{rev}$ oscillatory hub shears and moments. For a four bladed rotor, the objective function is given by

$$J = K_S \sqrt{(F_{4x})^2 + (F_{4y})^2 + (F_{4z})^2} + K_M \sqrt{(M_{4x})^2 + (M_{4y})^2 + (M_{4z})^2}$$

(1)

where $K_S$ and $K_M$ are appropriately selected weighting factors.

The vector of design variables $D$ consists of the thicknesses $t_1$, $t_2$, $t_3$, and the non-structural mass $m_{ns}$ located at the shear center, as shown in Fig. 3. The design variables are specified at several spanwise locations, and have side constraints to prevent them from reaching impractical values; these are stated as

$$D_j^{(L)} \leq D \leq D_j^{(U)}, \quad j = 1, 2, ..., N_{dv}.$$  

(2)

Figure 3. Simplified model of the blade structural member.
In addition, three types of behavior constraints, given by
\[ g_i(D) \leq 0, \quad i = 1, 2, ..., N_c, \tag{3} \]
are placed on the design variables. The first type of behavior constraints are frequency placement constraints, which are prescribed upper and lower bounds on the fundamental flap, lag, and torsional frequencies of the blade. The frequency placement constraints on the fundamental flap frequency are written as
\[ g_{\text{flap}}(D) = \frac{\omega_{F1}}{\omega_U} - 1 \leq 0 \tag{4} \]
and
\[ g_{\text{flap}}(D) = 1 - \frac{\omega_{F1}}{\omega_L} \leq 0 \tag{5} \]
where \( \omega_U \) and \( \omega_L \) are the prescribed upper and lower bounds on the fundamental flap frequency. Similar constraints are placed on the lag and torsional frequencies, i.e. \( g_{\text{lag}} \) and \( g_{\text{torsion}} \). In addition, all blade frequencies must differ from integer multiples of the angular velocity \(-1/\text{rev}, 2/\text{rev}, 3/\text{rev}, ... \), etc. – to avoid undesirable resonances.

Another behavior constraint is an autorotational constraint, which ensures that mass redistributions produced during the optimization do not degrade the autorotational properties of the rotor. Although there are several indices which can be used to represent the autorotational properties of the blade, the one used in this study is to require that the mass polar moment of inertia of the rotor be at least 90% of its baseline value. Mathematically, this is expressed as
\[ g(D) = 1 - \frac{J_P}{0.9J_{P0}} \leq 0 \tag{6} \]
where \( J_P \) is the mass polar moment of inertia of the rotor when it is spinning about the shaft, and \( J_{P0} \) is the baseline value.

The last behavior constraints are aeroelastic stability margin constraints, expressed mathematically as
\[ g_k(D) = \zeta_k + (\zeta_k)_{min} \leq 0, \quad k = 1, 2, ..., N_m \tag{7} \]
where \( N_m \) is the number of normal modes, \( \zeta_k \) is the real part of the hover eigenvalue for the \( k \)th mode, and \( (\zeta_k)_{min} \) is the minimum acceptable damping level for the \( k \)th mode. It should be noted that the most critical modes for stability are usually the first and second lag modes.

### IV. Efficient Global Optimization Algorithm

The goal in using global approximation, or surrogate, methods is to replace the true objective function with smooth functional relationships of acceptable accuracy that can be evaluated quickly. In order to construct the surrogate, the objective function must first be evaluated over a set of design points. The surrogate is then generated by fitting the initial design points. Although function evaluations, which come from the expensive helicopter simulations, are needed to form the approximation, the initial investment of computer time associated with function evaluations is significantly reduced compared to non-surrogate based optimization methods.

In this study, the objective function given by Eq. 1 is replaced by a surrogate function. This is different than the approach used in Ref. 16, where surrogates were fit to the six hub shears and moments and then used to form the objective function. In Ref. 16, the focus of the study was on the ability of approximation methods to capture the behavior of the vibratory loads, whereas in this study the focus is on an improved algorithm for searching the design space. Therefore, for simplicity, a single surrogate is created to replace the objective function.

Once the surrogate objective function is obtained, an obvious method for finding the optimum is to optimize the surrogate directly, i.e. the “one-shot” approach. However if the surrogate is not accurate everywhere in the design space, the optimization may lead to a local optima. Therefore, it is desirable to account for the uncertainty in the surrogate model since promising designs could lie in regions where the surrogate is inaccurate. The Efficient Global Optimization (EGO) algorithm proposed by Jones et al. is an alternative to the “one-shot” approach which accounts for uncertainty in the surrogate and is more
efficient. In EGO, a small number of initial design points are used to fit a kriging approximation. Based on the stochastic nature of kriging, an expected improvement function (EIF) is created in order to facilitate the selection of additional sample points (infill samples) where expensive computer simulations are to be conducted. These sample points are chosen to be where there is a high probability of producing a superior design over the current best design and/or where the predictions of the surrogate are unreliable due to a high amount of uncertainty. These infill samples represent a balance between the local consideration of finding an optimal design based on the information in the surrogate, and the global consideration of sampling in the design space where there is much uncertainty in the surrogate’s predictions. The kriging model is refit after the additional sample data is added to the initial data set, and the process of choosing additional sample points is repeated until a user defined stopping criterion is reached. Possible stopping criteria could be a predefined amount of computer time, when the expected improvement is low, or when the current best design is not significantly better than the best design from the previous iteration. The EGO algorithm is summarized in Fig. 4.

A. Design of Computer Experiments

When the initial data set is produced by a deterministic computer code (as is the case in the vibration reduction problem), the term “design of computer experiments,” is more appropriate than design of experiments.\textsuperscript{11, 29} The distinction is necessary because in physical experiments there is measurement error and other random sources of noise that cannot be controlled, and this affects the choice of the design point. However, in computer experiments, there is no random error; i.e., for a deterministic computer code, a given input will always yield the same output. Thus, the design of computer experiments need only be space-filling. Figure 5 illustrates the difference between a conventional design of experiment and a space-filling design. In the figure, locations of design points where experiments are to be conducted, which in this case represent design points where aeroelastic response simulations are performed, are illustrated for a design space which

---

**Figure 4. Efficient Global Optimization (EGO) algorithm**

The advantages of such a method over the “one-shot” approach are (1) a more global search is conducted by sampling in regions with high uncertainty in the surrogate, and (2) fewer expensive function evaluations are required since a smaller initial sample set is used and additional sample points are selected in a more “intelligent” manner, as opposed to starting with a larger initial data set. Details on how the initial data set is created, the kriging formulation, and the EIF formulation are given below.
has two design variables.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{Design of physical experiment vs. design of computer experiment}
\end{figure}

A commonly used space-filling design is Latin hypercube sampling (LHS).\textsuperscript{30} In LHS, each design variable is partitioned into \( N_{sp} \) equally spaced sections, or strata. Every design variable \( D_i \), where \( i = 1, 2, \ldots, N_{dv} \), is sampled once in each strata, which forms \( N_{dv} \) vectors of size \( N_{sp} \). The components of the \( N_{dv} \) vectors are then randomly combined to form an \( N_{sp} \times N_{dv} \) matrix known as a Latin hypercube, where each row corresponds to a design point at which a computer experiment is performed. A major disadvantage of Latin hypercube sampling is that design points can cluster together due to the random process by which design points are created. To prevent this, Optimal Latin hypercube (OLH)\textsuperscript{12} sampling is used in this study to ensure a more uniform (or space-filling) design of computer experiment. Optimal Latin hypercube sampling creates a more uniform design than conventional LHS by maximizing a spreading criterion, rather than randomly creating design points from the samples. Details on the OLH algorithm used in the present study can be found in Ref. 31. Figure 6 illustrates the difference between a conventional Latin hypercube and an optimal Latin hypercube.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{Conventional LH vs. optimal LH in two dimensional design space.}
\end{figure}

The points in the OLH represent design points at which helicopter simulations are to be conducted. Each simulation can be run independently of simulations at other design points. Therefore, the initial set of design points is generated using parallel computation in this study.

\section{Kriging}

Once an initial set of fitting points have been produced, it is possible to create the surrogate function. Kriging is based on the fundamental assumption that errors are correlated. This implies that one assumes the errors at two points close together will be close. In fact, the assumption that the errors are uncorrelated is only appropriate when the sources of error are random, such as in the case of measurement error or noise.
In the case of deterministic computer simulations, there is no source of random error. Therefore, it is more reasonable to assume that the error terms will be correlated and that this correlation is higher the closer two points are to each other.

Suppose an unknown function of \( N_{dv} \) design variables needs to be approximated and has been evaluated at \( N_{sp} \), sample points. Sample point \( i \) is denoted \( \mathbf{x}^{(i)} = (x_{1}^{(i)}, ..., x_{N_{dv}}^{(i)}) \). In kriging, the unknown function \( y(\mathbf{x}) \) is assumed to be of the form

\[
y(\mathbf{x}) = f(\mathbf{x}) + Z(\mathbf{x})
\]

where \( f(\mathbf{x}) \) is an assumed function (usually polynomial form) and \( Z(\mathbf{x}) \) is a realization of a stochastic (random) process which is assumed to be a Gaussian process with zero mean and variance of \( \sigma_{var}^{2} \) (i.e. \( Z(\mathbf{x}) \) follows a normal, or Gaussian, distribution). The function \( f(\mathbf{x}) \) can be thought of as a global approximation of \( y(\mathbf{x}) \), while \( Z(\mathbf{x}) \) accounts for local deviations which ensure that the kriging model interpolates the sample points exactly. The covariance matrix of \( Z(\mathbf{x}) \), which is a measure of how strongly correlated two points are, is given by

\[
\text{Cov}[Z(\mathbf{x}^{(i)}), Z(\mathbf{x}^{(j)})] = \sigma_{var}^{2} \mathbf{R}_{krg}
\]

where each element of the \( N_{sp} \times N_{sp} \) correlation matrix \( \mathbf{R}_{krg} \) is given by

\[
(R_{krg})_{ij} = R_{krg}(\mathbf{x}^{(i)}, \mathbf{x}^{(j)})
\]

and \( R_{krg}(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) \) is a correlation function which accounts for the effect of each interpolation point on every other interpolation point. This function is called the spatial correlation function (SCF) and is chosen by the user. The most commonly used SCF is the Gaussian correlation function,

\[
R_{krg}(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) = \exp \left[ - \sum_{k=1}^{N_{dv}} \vartheta_{k} |x_{k}^{(i)} - x_{k}^{(j)}|^{p_{k}} \right],
\]

which is also employed in this study. The Gaussian SCF is dependent on the distance between two points. As two points move closer to each other, \(|x_{k}^{(i)} - x_{k}^{(j)}| \rightarrow 0\) and Eq. 11 approaches unity, which is the maximum value of the Gaussian SCF. In other words, the Gaussian SCF recovers the intuitive property that the closer two points are to each other, the higher the correlation between the points.

The fitting parameters \( \vartheta_{k} \) and \( p_{k} \) are unknown correlation parameters which need to be determined. In order to determine these parameters, the form of \( f(\mathbf{x}) \) needs to be selected. The most common choice for \( f(\mathbf{x}) = f(\mathbf{x}) = \beta \) where \( \beta \) is a constant. Previous studies have found that modeling with the SCF is so effective, that using a constant for the global behavior results in little loss of fidelity.\(^{14,29,32,33}\) Another common simplification, is to fix all \( p_{k} = 2 \). When this simplification is combined with the constant global approximation, the approximation method is known as \textit{ordinary kriging}. In Ref. 16, allowing \( p_{k} \) to vary during the fitting process did not offer any advantages over fixing all \( p_{k} = 2 \), therefore ordinary kriging is employed in this study.

In order to find \( \vartheta_{k} \) and \( p_{k} \), the generalized least square estimates of \( \beta \) and \( \sigma_{var}^{2} \), denoted by \( \hat{\beta} \) and \( \hat{\sigma}_{var}^{2} \), respectively, are employed:\(^{14,32}\)

\[
\hat{\beta} = (\mathbf{1}^{T} \mathbf{(R}_{krg}^{-1} \mathbf{1})^{-1}} \mathbf{1}^{T} \mathbf{(R}_{krg}^{-1} \mathbf{y}
\]

and

\[
\hat{\sigma}_{var}^{2} = \frac{\mathbf{(y - \mathbf{1}\hat{\beta})}^{T} \mathbf{(R}_{krg}^{-1} \mathbf{(y - \mathbf{1}\hat{\beta})}}}{N_{sp}^{-}}
\]

where \( \mathbf{1} \) is a vector populated by ones and \( \mathbf{y} \) is a vector of observed function outputs at the interpolation points; both vectors are of length \( N_{sp} \). With \( \hat{\sigma}_{var}^{2} \) and \( \hat{\beta} \) known, \( \vartheta_{k} \) and \( p_{k} \) are found such that a likelihood function\(^{14,32,34,35}\) is maximized. The likelihood function, given in Eq. 14, is a measure of the probability of the sample data (or interpolation data) being drawn from a probability density function associated with a Gaussian process. Since the stochastic process associated with kriging has been assumed to be a Gaussian process, one seeks the set of \( \vartheta_{k} \) and \( p_{k} \) that maximize the probability that the interpolation points have been drawn from a Gaussian process.

\[
- \frac{N_{sp}}{2} \ln(\hat{\sigma}_{var}^{2} + \ln |\mathbf{R}_{krg}|)
\]

The maximum likelihood estimates (MLE’s) of \( \vartheta_{k} \) and \( p_{k} \) represent the “best guesses” of the fitting parameters. Any values of \( \vartheta_{k} \) and \( p_{k} \) would result in a surrogate which interpolates the sample points exactly,
but the “best” kriging surrogate is found by optimizing the likelihood function. This auxiliary optimization process can result in significant fitting time depending on the size of the system. Due to the optimization process needed to create the kriging surrogate, kriging is only appropriate when the time needed to generate the interpolation points is much larger than the time to interpolate the data – which is the case in the helicopter vibration problem. With all parameters known, the kriging approximation to a function \( y(x) \) can be written as\(^{14, 29, 32, 33}\)

\[
\hat{y}_{\text{kr}g} = \hat{\beta} + r_{\text{kr}g}(x)^T (R_{\text{kr}g})^{-1} (y - 1) \hat{\beta}
\]

(15)

where

\[
r_{\text{kr}g}(x) = \begin{bmatrix} R_{\text{kr}g}(x, x^{(1)}) \, R_{\text{kr}g}(x, x^{(2)}) \, \ldots \, R_{\text{kr}g}(x, x^{(N_{\text{sp}})}) \end{bmatrix}^T
\]

(16)

The column vector \( r_{\text{kr}g}(x) \) of length \( N_{\text{sp}} \) is the correlation vector between an arbitrary point \( x \) and the interpolation points, \( x^{(1)}, \ldots, x^{(N_{\text{sp}})} \). In this study, the kriging surrogates were created using a MATLAB kriging toolbox.\(^{36}\)

C. Expected Improvement Based Search of the Design Space

Before forming the expected improvement function (EIF) it is necessary to derive an estimation of the error, or uncertainty, in the kriging surrogate can be obtained when \( I = 0 \). In the design space of the kriging predictor can be written as\(^{14}\)

\[
\text{MSE}(x) \equiv s^2(x) = \hat{\sigma}_{\text{var}}^2 \left[ 1 - r_{\text{kr}g}^T R_{\text{kr}g}^{-1} r_{\text{kr}g} + \frac{(1 - 1^T R_{\text{kr}g}^{-1} r_{\text{kr}g})^2}{1^T R_{\text{kr}g}^{-1} 1} \right].
\]

(17)

Equation 17 represents an approximation of the error in the kriging prediction at any design point. The advantage of Eq. 17 is that an estimate of the error, or uncertainty, in the kriging surrogate can be obtained without additional expensive function evaluations. As shown in Ref. 14, Eq. 17 is equal to zero when evaluated at one of the sample points used to create the kriging model. This is expected since the kriging approximation interpolates the sample points exactly, thus there is no uncertainty at a sample point.

With the kriging model and the MSE defined, the expected improvement function can be derived. For a minimization problem, the improvement over the current best design is written as\(^{18, 34}\)

\[
I = \max(y_{\text{min}} - \hat{y}(x), 0)
\]

(18)

where \( y_{\text{min}} \) is the best design out of all of the sample points used to create the surrogate. The expected improvement is simply the expected value of Eq. 18,\(^{34}\)

\[
E(I) \equiv EIF(x) = \begin{cases} 
(y_{\text{min}} - \hat{y}_{\text{kr}g}) \Phi_{\text{dist}} \left( \frac{y_{\text{min}} - \hat{y}_{\text{kr}g}}{s} \right) + s \phi_{\text{den}} \left( \frac{y_{\text{min}} - \hat{y}_{\text{kr}g}}{s} \right) & \text{if } s > 0 \\
0 & \text{if } s = 0
\end{cases}
\]

(19)

where \( \Phi_{\text{dist}}(*) \) is the standard normal distribution function and \( \phi_{\text{den}}(*) \) is the standard normal density function.

The first term in the EIF is the difference between the current best objective function value and the predicted response at an arbitrary design, \( x \), multiplied by the probability that \( y(x) \) is better than \( y_{\text{min}} \). This term is large where \( \hat{y}_{\text{kr}g} \) is likely to be better than \( y_{\text{min}} \). Similarly, the second term is large where the error metric \( s(x) \) is large and thus where there is much uncertainty in the surrogate’s prediction. The design points with the highest expected improvement represent a balance between finding promising regions of the design space based on information from the surrogate (local search) and finding regions of high uncertainty in the surrogate (global search). In EGO, the expected improvement function is maximized to find the set of design points with the highest expected improvement, and then additional simulations at these designs are conducted. In this study, a number of local optima equal to the number of available processors are found in order to take advantage of parallel computation. Similar to the applicability of parallel computing in the design of computer experiment phase, simulations at each local optima of the EIF can be run on individual processors.
It is possible to control the balance between the local and global search characteristics of the EIF by weighting the two terms. The weighted expected improvement function (WEIF) is given as

\[
WEIF(x) = \begin{cases} 
  w(y_{\text{min}} - \hat{y}_{\text{kr}}) & \text{if } s > 0 \\
  0 & \text{if } s = 0 
\end{cases}
\]

where \(0 \leq w \leq 1\). Setting \(w = 0\) would shift the balance towards the global search term in the EIF, while \(w = 1\) would result in a more local search. A value of \(w = 0.5\) results in the same characteristics of the balanced search based on the EIF. The effect of varying \(w\) on the performance of Efficient Global Optimization is investigated in this paper.

V. Results

This section presents optimization results using the Efficient Global Optimization algorithm. Parameters necessary for the helicopter simulation are given first, followed by optimization results.

A. Helicopter Parameters

The helicopter configuration used in all computations is given in Table 1. The simulations are conducted at an advance ratio of 0.15 and descent angle of 6°. These parameters are characteristic of a flight condition in which high vibration levels due to strong blade vortex interaction (BVI) are encountered.\(^{37}\) Figure 7 illustrates a helicopter in descent; this figure is also employed for the propulsive trim calculation.

<table>
<thead>
<tr>
<th>Dimensional Data</th>
<th>Non-Dimensional Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R = 4.91) m</td>
<td>(\Omega = 425) RPM</td>
</tr>
<tr>
<td>(N_b = 4)</td>
<td>(c = 0.05498 R)</td>
</tr>
<tr>
<td>(\beta_p = 0.0^\circ)</td>
<td>(C_{do} = 0.01)</td>
</tr>
<tr>
<td>(\theta_{pt} = 0^\circ)</td>
<td>(\alpha_d = 6^\circ)</td>
</tr>
<tr>
<td>(\mu = 0.15)</td>
<td>(C_W = 0.005)</td>
</tr>
<tr>
<td>(\sigma = 0.07)</td>
<td>(C_{df} = 0.01)</td>
</tr>
<tr>
<td>(X_{FA} = 0.0)</td>
<td>(Z_{FA} = 0.3)</td>
</tr>
<tr>
<td>(X_{FC} = 0.0)</td>
<td>(Z_{FC} = 0.3)</td>
</tr>
</tbody>
</table>

The weighting factors in the objective function, \(K_S\) and \(K_M\), are selected to be 1. These weighting factors result in an objective function which represents the sum of the 4/rev oscillatory hub shear resultant and the 4/rev oscillatory hub moment resultant in the hub-fixed non-rotating frame. For this study, the following side constraints are enforced:

\[
1.0\, \text{mm} \leq t_1 \leq 8.0\, \text{mm} \tag{21}
\]

\[
1.0\, \text{mm} \leq t_2, t_3 \leq 12.0\, \text{mm} \tag{22}
\]

\[
0.0 \leq m_{ns}/m_0 \leq 0.25 \tag{23}
\]

The upper and lower bounds used for the frequency placement constraints are given in Table 2, and are similar to those used in Ref. 38, which also used cross-sectional dimensions as design variables.

In the aeroelastic stability constraints given by Eq. 7, the minimum acceptable damping for all modes, \((\zeta_k)_{\text{min}}\), is chosen to be 0.01, as in Ref. 8. Additionally, the constraints are modified for the 2\(^{nd}\) lag mode, which can sometimes be slightly unstable. To prevent this situation, a small amount of structural damping is added to this mode, as in Ref. 2. For this study, 0.5% structural damping is added to stabilize the 2\(^{nd}\) lag mode of the baseline blade.
Figure 7. Helicopter in descent flight condition.

Table 2. Upper and lower bounds on the fundamental frequencies (/rev).

<table>
<thead>
<tr>
<th></th>
<th>Flap</th>
<th>Lag</th>
<th>Torsion</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega_U$</td>
<td>1.20</td>
<td>0.80</td>
<td>6.50</td>
</tr>
<tr>
<td>$\omega_L$</td>
<td>1.05</td>
<td>0.60</td>
<td>2.50</td>
</tr>
</tbody>
</table>

The rotor blade was discretized into the 6 finite elements shown in Fig. 8 and the design variables were defined at two nodal locations – the 68% and tip stations – resulting in a total of 8 design variables. These two blade stations were chosen because previous studies have shown that (a) non-structural masses are most effective for vibration reduction when they are distributed over the outboard 1/3 of the blade\textsuperscript{39,40} and (b) a similar configuration has been used successfully for vibration reduction in Ref. 28. The root cross-sectional variables were fixed at baseline values\textsuperscript{16} and the cross-sectional variables were assumed to vary linearly between spanwise stations. The non-structural mass at the elastic axis inboard of the 68% station was set to zero.

Figure 8. Finite element discretization of the blade.
B. Optimization Results

In this section, optimization results using EGO with a weighted expected improvement function are compared to “one-shot” optimization results. To start the EGO algorithm, an optimal Latin hypercube (OLH) of 100 design points was chosen. Since the computational cost of evaluating the actual constraints is low, the constraints at each point in the OLH were evaluated first. Designs that violated the constraints were removed from the data set so that expensive forward flight simulations would not be wasted. From the 100 point OLH, 53 blade designs were feasible and had converged trim solutions. In addition, OLH’s of 250 and 500 points, resulting in 143 and 284 feasible designs respectively, were used to obtain “one-shot” optimization results.

For each iteration of the EGO algorithm (see Fig. 4), 25 local optima of the WEIF were selected for parallel computation. In addition, a variation of the EGO algorithm was considered in which the 25 sample points for each iteration were given by the local optima from direct optimization of the surrogate objective function using a genetic algorithm; these results are referred to as “GA updating”. In order to ensure that each of the 25 designs were different blade designs, the following criterion was used: a design was considered to be different from another design if there was at least a 0.5 mm difference in any of the thickness design variables or if there was a difference of at least 5% of the root non-structural mass in any of the non-structural mass design variables. Each simulation took about 6 hours to complete on a Linux cluster with 1.8 GHz Opteron processors. The EGO algorithm was run for at least 2 iterations and was stopped if the best design from the current iteration was not better than the best blade design from previous iterations. Optimization of the WEIF and the surrogate objective function was conducted with the Multi-Island Genetic Algorithm in the iSIGHT software package. All optimization results are compared to a baseline blade with cross-sectional properties resembling an MBB BO-105 blade, which is considered to have good vibration characteristics.

Table 3 shows the results for “one-shot” optimization of kriging surrogates built from optimal Latin hypercubes with 53, 143 and 284 design points, compared to the best design in each OLH. For all three kriging models in Table 3, the optimal design returned by the genetic algorithm optimization is worse than the best design in the optimal Latin hypercube used to create the surrogate. This is because at certain designs, the kriging surrogates predict a negative value for the vibration objective function, which is not realistic given Eq. 1. So when the surrogates are minimized directly, the optimization algorithm will move toward regions with negative objective function values, even though the surrogates are inaccurate in these regions. This illustrates the danger of directly optimizing a surrogate for a function with behavior that is difficult to capture: if the surrogate is not accurate everywhere in the design space, direct optimization may lead to a region in the design space with inferior designs.

<table>
<thead>
<tr>
<th># of feasible designs from OLH</th>
<th>Best % Vibration Reduction from the OLH</th>
<th>Actual % Vibration Reduction at the “One-Shot” Optimal Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>53</td>
<td>18.52 %</td>
<td>4.50 %</td>
</tr>
<tr>
<td>143</td>
<td>41.60 %</td>
<td>17.47 %</td>
</tr>
<tr>
<td>284</td>
<td>53.39 %</td>
<td>33.58%</td>
</tr>
</tbody>
</table>

Efficient Global Optimization results based on updating the surrogate model with additional blade designs selected by the WEIF and the genetic algorithm are shown in Table 4 and Fig. 9. Table 4 summarizes the amount of vibration reduction achieved, while Fig. 9 shows the best value of the vibration objective function after each iteration of the EGO algorithm.

The results for $w = 0.5$ would also be the results if an unweighted expected improvement function were considered. For $w = 0.5$, the maximum amount of vibration reduction is 54.55%, and the algorithm finishes after 153 objective function evaluations. This is clearly superior to the 53.39% vibration reduction obtained with 284 objective function evaluations. This comparison indicates that the expected improvement function is effective in reducing the number of function evaluations needed to find optimal designs compared to using a relatively large number of sample points from a space-filling scheme. Other than $w = 0.4$ and $w = 0.6$, every weight setting resulted in vibration reduction greater than 53.39%, but with less function evaluations than the 284 in Table 3. Therefore, other than at $w = 0.4$ and $w = 0.6$, the WEIF was superior to the
Table 4. Vibration Reduction (relative to MBB BO-105 baseline blade) using EGO.

<table>
<thead>
<tr>
<th>EGO Setting</th>
<th>Number of Function Evaluations</th>
<th>% Vibration Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>( w = 0.0 )</td>
<td>72</td>
<td>55.15 %</td>
</tr>
<tr>
<td>( w = 0.2 )</td>
<td>128</td>
<td>65.18 %</td>
</tr>
<tr>
<td>( w = 0.4 )</td>
<td>128</td>
<td>45.00 %</td>
</tr>
<tr>
<td>( w = 0.5 )</td>
<td>153</td>
<td>54.55 %</td>
</tr>
<tr>
<td>( w = 0.6 )</td>
<td>128</td>
<td>44.09 %</td>
</tr>
<tr>
<td>( w = 0.8 )</td>
<td>178</td>
<td>64.53 %</td>
</tr>
<tr>
<td>( w = 1.0 )</td>
<td>153</td>
<td>54.35 %</td>
</tr>
<tr>
<td>GA updating</td>
<td>178</td>
<td>54.68 %</td>
</tr>
</tbody>
</table>

“one-shot” approach. Furthermore, settings of \( w = 0.2 \) and \( w = 0.8 \) resulted in the best designs with 65.18% and 64.53% reduction respectively. The problem considered in this study is an example of a situation in which a promising design was located in a region of high uncertainty, and with a more global setting of \( w = 0.2 \), the EGO algorithm was able to locate it. Note that in the results for \( w = 0 \), the surrogate is not updated with 25 additional sample points each iteration because only 10 of the local optima of the WEIF function from the first iteration were feasible and had converged trim solutions, while only 9 designs were feasible with converged trim solutions from the second iteration. Thus for \( w = 0 \), a total of 72 objective function evaluations were computed.

The importance of accounting for uncertainty is further illustrated by comparing the results in Fig. 9 for \( w = 0.8 \) and 1.0 with the results corresponding to genetic algorithm updating. For \( w = 0.8 \), a superior design is found compared to the best design from genetic algorithm updating. Thus even a small weighting on the global search characteristics was beneficial compared to a purely local search. The setting of \( w = 1.0 \) results in essentially the same amount of vibration reduction as the genetic algorithm updating, but in 1 less
iteration of EGO. Therefore, the \( w = 1.0 \) setting resulted in a more efficient optimization algorithm than genetic algorithm updating. This can be explained by examining the normal distribution term (which is an explicit function of the uncertainty measure \( s(x) \)) in Eq. 20, since this term accounts for the only difference between optimization of the WEIF with \( w = 1.0 \) and direct optimization of the surrogate objective function \( \hat{y}_{\text{kr}} \). For an amount of predicted reduction given by \( y_{\text{min}} - \hat{y}_{\text{kr}} \), the normal distribution function will be high for low values of \( s(x) \) and vice versa. Therefore, maximization of \( (y_{\text{min}} - \hat{y}_{\text{kr}})\Phi_{\text{dist}}(\frac{y_{\text{min}} - \hat{y}_{\text{kr}}}{s(x)}) \) will favor blade designs which are predicted to lower vibrations and at which the uncertainty in the surrogate is low. In contrast, direct optimization of the surrogate objective function using the genetic algorithm will favor designs which are predicted to lower vibration, but the uncertainty in the surrogate’s predictions is not explicitly accounted for. Thus the inclusion of the uncertainty measure during optimization accounts for the improved efficiency of the EGO algorithm when using a setting of \( w = 1.0 \) compared to the variation of EGO based on genetic algorithm updating.

Figures 10 a-h show the design variables corresponding to the best designs as EGO progresses. These results show that for this problem, the progression of design variables during EGO is different for each weight setting, and thus the performance of EGO is sensitive to the weight setting. Even though all weight settings resulted in significant vibration reduction, each setting led to a different optimal design. This suggests that there are numerous local optima in the vibration objective function at this flight condition.

Figures 11 a-f show the variation of the vibratory shears and moments corresponding to the best designs as the EGO algorithm progresses. Every weight setting resulted in designs with significant reduction (over 70%) in the vertical shear \( F_{4Z} \), while other vibratory components may increase depending on the weight setting. Figure 12 compares the baseline hubloads with those corresponding to the best design from EGO. In Fig. 12 it is clear that the vertical shear is the dominant vibratory component of the baseline blade, and therefore the majority of the reduction in the objective function comes from the significant reduction in the vertical shear.

The errors in the surrogates are given after each iteration of EGO in Fig. 13. The errors were calculated by comparing the predicted response of the surrogates to the actual response at 427 test points, which were obtained by combining the sample data of the 143 point and 284 point OLH’s from Table 3. The errors in the surrogates tend to increase as additional sample points were added during the optimization. This is because only regions of high expected improvement are sampled in EGO, which led to a decline in the predictive capabilities of the surrogate in regions of low expected improvement. Although the accuracies diminished slightly, the more global weightings of \( w = 0.0 \) and \( w = 0.2 \) had negligible effects on the accuracy of the surrogates, which is not surprising since the goal of using more global weightings is to reduce the uncertainty, or error, in the surrogate.

VI. Conclusions

The results in this paper demonstrate that the expected improvement function can be used effectively to reduce the number of objective function evaluations needed to achieve significant levels of vibration reduction. Beginning with a relatively small number of initial fitting points and using the weighted expected improvement function to determine additional designs at which to conduct helicopter simulations proved to be more efficient than starting with a larger number of space-filling sample points. The principle results from this study are summarized below.

1. Using the weighted expected improvement function, vibration reduction of 53-65% was achieved with up to 178 objective function evaluations, while 53% vibration reduction was achieved with 284 function evaluations when simply using a large number of space-filling sample points.

2. The results illustrate that it is important to account for the uncertainty in the surrogate during optimization. Otherwise, promising designs may be missed and efficiency may be sacrificed.

3. Although the surrogates were not accurate everywhere in the design space, updating the surrogates in parallel with an Efficient Global Optimization algorithm was effective.
Figure 10. Progression of the best blade design during EGO.
Figure 11. Progression of the vibratory loads at the best designs during EGO.
Figure 12. Comparison of the vibratory loads of the best design from EGO and the baseline blade.

Figure 13. Errors in the surrogates after each iteration of EGO.
VII. Acknowledgments

This research was supported in part by the FXB Center for Rotary and Fixed Wing Air Vehicle Design, and by ARO grant 02-1-0202 with Dr. B. LaMattina as grant monitor.

References


