

A SURROGATE APPROACH TO THE EXPERIMENTAL OPTIMIZATION OF MULTIELEMENT AIRFOILS

John C. Otto*

Multidisciplinary Optimization Branch
 NASA Langley Research Center
 Hampton, Virginia
 Department of Aeronautics and Astronautics, M.I.T.

Drew Landman†

Department of Engineering Technology
 Old Dominion University
 Norfolk, Virginia

Anthony T. Patera‡

Department of Mechanical Engineering
 Massachusetts Institute of Technology
 Cambridge, Massachusetts

Abstract

The incorporation of experimental test data into the optimization process is accomplished through the use of Bayesian-validated surrogates. In the surrogate approach, a surrogate for the experiment (e.g., a response surface) serves in the optimization process. The validation step of the framework provides a qualitative assessment of the surrogate quality, and bounds the surrogate-for-experiment error on designs “near” surrogate-predicted optimal designs. The utility of the framework is demonstrated through its application to the experimental selection of the trailing edge flap position to achieve a design lift coefficient for a three-element airfoil.

Introduction

To address the inherent difficulties in examining many design points experimentally, a three-element airfoil model with internally embedded actuators has been developed.¹ The model (Fig. 1) has a nested chord of $c = 18$ in., a span of $b = 36$ in., and was designed for low-speed testing in several local tunnels,

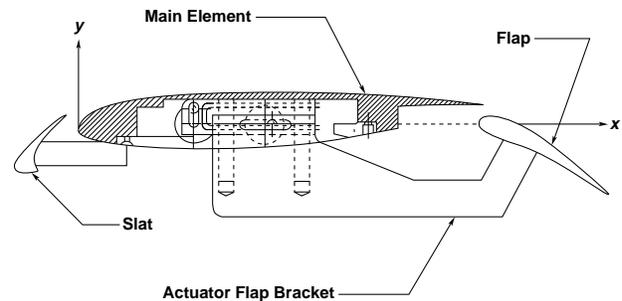


Figure 1: Three-element model with internal flap actuators.

including the NASA Langley Research Center 2- by 4-foot and the Old Dominion University (ODU) 3- by 4-foot low-speed facilities. The main element chord is $c_{main} = 14.95$ in., and the flap and slat chords (expressed as a percentage of the nested chord) are 30 and 14.5 percent, respectively. The flap and slat are both deflected to 30° for all tests. Although this particular model is suitable only for low Reynolds number testing, the techniques developed should be applicable to higher Reynolds number testing as well.

The flap actuators are computer controlled and position the flap horizontally and vertically (x and y , respectively). The model has been used in the ODU tunnel to compile baseline values for lift coefficient C_l versus flap gap and overhang at fixed angles of attack and slat riggings. A first-order optimizer that uses a variant of the method of steepest ascent^{2,3} has been demonstrated in real time.⁴ The capability of the computer controller to automatically take data

*Research Scientist, Member AIAA

†Associate Professor, Member AIAA

‡Professor

Copyright © by the American Institute of Aeronautics and Astronautics, Inc. No copyright is asserted in the United States under Title 17, U.S. Code. The U.S. Government has a royalty-free license to exercise all rights under the copyright claimed herein for Governmental Purposes. All other rights are reserved by the copyright owner.

at a prescribed set of (x, y) coordinates makes this setup ideal for the surrogate methods described next.

The Bayesian-validated surrogate framework applied in this paper provides a practical means to incorporate experimental data directly into the design optimization process. In the surrogate approach to optimization, a surrogate (i.e., a simplified model, for example a response surface) for the experiment is constructed from *off-line* appeals to the experiment. The surrogate is then used in subsequent optimization studies. This approach to optimization can be contrasted with *on-line* (direct insertion) strategies, in which appeals to the experiment are embedded directly into the optimization process.

The off-line surrogate approach^{5–8} to optimization offers several advantages to on-line approaches. First, by construction, surrogates are computationally inexpensive and are thus easily incorporated into optimization procedures. Additionally, the low computational requirements create a highly interactive and flexible design environment, which allows the designer to easily pursue and examine multiple design points. Second, the number of appeals to the experiment or simulation is known *a priori*, which ensures that the design can be accomplished without exhausting available resources. Third the surrogate approach offers a natural means to incorporate data from previous runs and/or other sources.

As regards disadvantages, the primary drawback is that in high dimensional design spaces, surrogate construction is difficult and design localization is poor. A second limiting factor in the application of the surrogate approach to experimental tests is the need to validate the surrogate at input points chosen randomly in the design space. This capability, present in the experiment central to this work, is not typical of most experimental tests. Finally, surrogate-based optimization introduces a new source of error. The surrogate validation strategy and error norms discussed in this paper seek to quantify the discrepancy between the surrogate and the experiment by providing estimates to the system predictability and optimality.

In this paper, we first describe the experimental model and the testing methods used. Second, we present the optimization problem that is central to the work. Third, we briefly describe the three steps of the baseline surrogate framework (i.e., construction/validation, surrogate-based optimization, and *a posteriori* error analysis), summarize the inputs to the framework, and then present an overview of the more sophisticated surrogate algorithms. Finally, we present sample results obtained from the surrogate framework for output maximization and multiple-target designs, and compare the surrogate approach

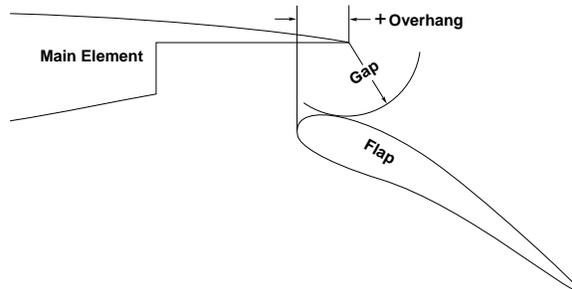


Figure 2: Definition of gap and overhang.

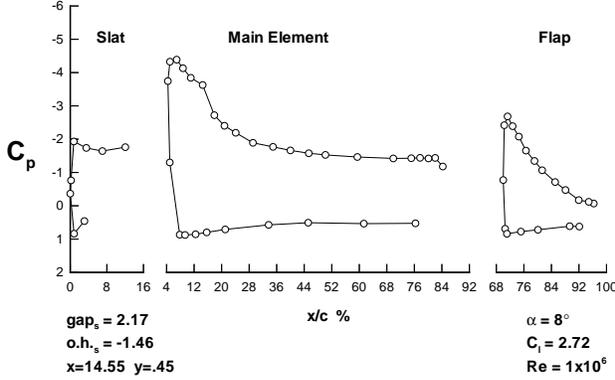
with the direct insertion results reported previously.⁴

Experimental Testing Methods

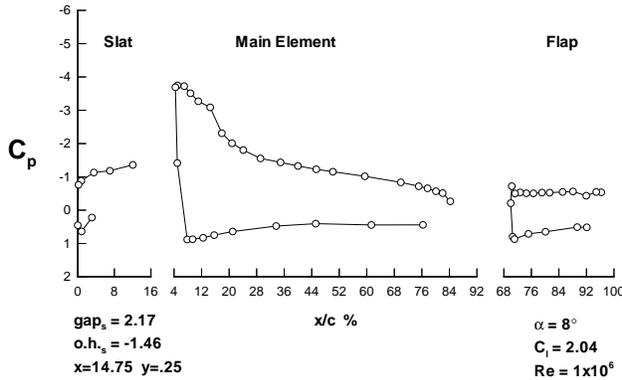
An important practical problem encountered in wind-tunnel testing of multielement airfoils is the need to test a range of configurations to ensure that the optimum is selected. Unfortunately, this testing can be prohibitively time consuming if one considers all possible variables, such as flap position and deflection, slat position and deflection, overall angle of attack, and Reynolds number. For example, a range of flap locations and orientations relative to the main element is typically tested. In a cryogenic or pressurized facility, model geometry changes necessitate large delays in testing. These delays often result in investigators choosing a sparse test matrix and an optimum that is based on only a few points. The ability to move the flap under computer control provides a unique opportunity to explore the entire range of useful gap and overhang values (Fig. 2).

In this experiment, the flap actuators, tunnel flow setting, and data acquisition were controlled by a personal computer running Lab View⁹ software. A program was written to allow any number of flap positions (in x and y) to be sampled in any order. Wind tunnel power was controlled such that at the beginning of each test the tunnel was restarted to avoid hysteresis effects.⁴ The experimental setup allowed the user to start the program, which at each location in turn automatically measured the free-stream properties, sampled and recorded pressures around the centerline of the model, and then calculated lift coefficients for the three-element airfoil. This process required approximately 2 min. for each data point.

Two typical pressure distributions are shown in Figure 3, where the ordinate is the pressure coefficient C_p and the abscissa is distance from the leading edge expressed as a percent of the nested chord. The data for Figure 3(a) represents a point near the peak C_l for this configuration, and the plot in Figure 3(b)



(a) Fully attached flow over all elements.



(b) Detached flow over the flap.

Figure 3: Experimental pressure data.

indicates full separation over the flap.

Test matrices were developed to survey flap positions, which ranged from approximately 0.8 – 3.5 percent (gap) and $-0.4 - 3.4$ percent (overhang) relative to the nested chord c . Two angles of attack and two slat geometries were selected. An angle of attack α of 8° was chosen as representative of an approach value. An α of 14° represented the limit of good-quality two-dimensional flow for the ODU tunnel installation without tunnel wall boundary-layer control. Two slat settings were chosen: a slat gap of 3.03 percent with an overhang of 2.46 percent and, for a smaller gap setting, a slat gap of 2.17 percent with a slat overhang of -1.46 percent

Positional accuracy was enhanced by requiring that the flap move to a reference point above and behind the desired evaluation points ($x_{ref} > x_{eval}$, $y_{ref} > y_{eval}$) and then back to the evaluation point. This eliminated any effect of backlash in the mechanical drive-train. Two simple tests provided an indication of the inherent collective error due to instrumentation and positioning. The first test involved two separate evaluation points; the first point was

in a region in which the flow was known to be fully attached to all elements, and the second point was chosen in a region in which flow over the flap was fully separated. The positioning program was used to move the flap between a reference point and one of the evaluation points. The tunnel was restarted before every evaluation, and the test was repeated 30 times in each case. The standard deviation of C_l was found to be 0.004 for the separated case (0.16 percent) and 0.0118 for the attached case (.36 percent). For the second test, the program automatically sampled 29 points over the entire test region for two different trials. The error in C_l between the two runs averaged 0.71 percent with a standard deviation of 0.75 percent. Although these tests are not exhaustive, they do provide a benchmark for the C_l error.

The turbulence intensity in the ODU tunnel was measured at less than 0.2 percent. Flow quality over the model was monitored through 12 spanwise taps: 6 on the flap, and 6 on the main element. The flow was considered to be two-dimensional if the magnitude of the spanwise nonuniformity was less than 5 percent of the total C_p variation over the entire model.¹⁰ The data presented are uncorrected for boundary effects were taken at a Reynolds Re number of 1×10^6 based on the nested chord.

Optimization Problem

We begin by introducing a vector \mathbf{p} of M design inputs that lie in the input (or “design”) domain $\Omega \subset \mathbb{R}^M$, an input-output function $\mathcal{S}(\mathbf{p}) : \Omega \rightarrow \mathbb{R}$, and an objective function $\Psi(\mathcal{S}(\mathbf{p}), \mathbf{p}, \lambda)$ that characterizes our design goals, where λ is a vector (or possibly scalar) design parameter. For the work presented here, we set $\mathbf{p} = (x, y)$ (the x - and y -positions of the flap) as the $M = 2$ inputs and restrict ourselves to an input domain Ω of reasonable flap positions (described in more detail in the results section). The output of interest is the lift coefficient, $\mathcal{S}(\mathbf{p}) = C_l(x, y)$. The objective function is $\Psi(\mathcal{S}(\mathbf{p}), \mathbf{p}, \lambda) = |\mathcal{S}(\mathbf{p}) - \lambda|$ which has been referred to as the “discrimination” problem.¹¹

With the above terms defined, the minimizer $\mathbf{p}^* = (x^*, y^*)$ to the exact optimization problem is given by

$$\mathbf{p}^* = \arg \min_{\mathbf{p} \in \Omega} |\mathcal{S}(\mathbf{p}) - \lambda|. \quad (1)$$

In this formulation, the goal is to find that (or “an”) input vector $\mathbf{p}^* = (x^*, y^*)$ that achieves as closely as possible the target lift coefficient value λ . If the target lift coefficient λ is set sufficiently small (large), the formulation describes the output minimization (maximization) problem, assuming that $\mathcal{S}(\mathbf{p})$ is bounded from below (above).

In the on-line approach, the experiment is invoked at every optimization step needed to solve Equation (1). In the off-line approach, a surrogate, $\tilde{\mathcal{S}}(\mathbf{p}) \approx \mathcal{S}(\mathbf{p})$, for the experiment is inserted into the optimization problem. The minimizer, $\tilde{\mathbf{p}}^* = (\tilde{x}^*, \tilde{y}^*)$, for the resulting, surrogate-based, discrimination problem is then given by

$$\tilde{\mathbf{p}}^* = \arg \min_{\mathbf{p} \in \Omega} |\tilde{\mathcal{S}}(\mathbf{p}) - \lambda|. \quad (2)$$

Here, the optimization proceeds exactly as it would for the on-line approach, but the lift coefficient surrogate $\tilde{\mathcal{S}}(\mathbf{p})$ is invoked instead of the experiment. The surrogate problem that corresponds to Equation (1), but with a general objective function $\Psi(\mathcal{S}(\mathbf{p}), \mathbf{p}, \lambda)$, has been reported by Yeşilyurt¹² and Yeşilyurt and Patera.⁸

Surrogate Framework

The advantages to pursuing a surrogate-based approach to optimization have already been described. However, to use a surrogate-based approach with confidence in a design setting, the issues of predictability and optimality must be addressed.¹³ For predictability, the concern is with how the actual experiment performs in the vicinity of the surrogate-predicted minimizer $\tilde{\mathbf{p}}^*$. If the surrogate-predicted minimizer is to be of value, we must be able to bound $|\mathcal{S}(\mathbf{p}') - \mathcal{S}(\tilde{\mathbf{p}}^*)|$ for \mathbf{p}' “near” $\tilde{\mathbf{p}}^*$, and this bound must be acceptably small. In the case of optimality, the designer requires confidence that the surrogate-predicted optimizer $\tilde{\mathbf{p}}^*$ is near the “exact” optimizer, that is, $\tilde{\mathbf{p}}^* \approx \mathbf{p}^*$. Optimality requires stronger assumptions in regard to the form of the objective function (e.g., quasi-convexity) and is, therefore, difficult to determine in real applications. Optimality is, however, an important consideration and, although not addressed further here, has been examined elsewhere.^{8,12}

The distinguishing attribute of the Bayesian-validated surrogate methodology is that a complete and rigorous validation step is fully integrated into the *a posteriori* error analysis of the surrogate-predicted design(s). The approach described here is related to probably-approximately-correct approaches^{14,15} and information-based complexity theory.¹⁶ The surrogate approach differs, however, from the former in that it is truly non-parametric (no assumption is made in regard to the distribution of $\tilde{\mathbf{p}}^*$) and from the latter in that it requires no regularity estimates for the input-output function.

The surrogate approach is broken into three steps. In the first stage, surrogate construction/validation,

experimental results and/or prior information are used to construct the approximation, $\tilde{\mathcal{S}}(\mathbf{p}) \approx \mathcal{S}(\mathbf{p})$; additional queries to the experiment are used to validate the approximation. In the second step of the process, surrogate-based optimization, solutions to surrogate optimization problem of Equation (2) are obtained. In the third and final step, *a posteriori* error analysis, the results of the validation are used to analyze the consequences of the surrogate-for-simulation substitution. In the following subsections, we describe the three steps of the *baseline* surrogate framework, summarize the inputs to the framework, and review the more sophisticated surrogate algorithms.

Construction/Validation

We construct the lift coefficient surrogate $\tilde{\mathcal{S}}(\mathbf{p}) = \mathcal{A}(\mathcal{X}^{co}) \approx \mathcal{S}(\mathbf{p})$ using an approximation scheme, $\mathcal{A}: (\mathbb{R}^M, \mathbb{R})^{N^{co}} \rightarrow L^\infty(\Omega)$ and a construction sample set of input-output pairs

$$\mathcal{X}^{co} = \{(\mathbf{p}_i, R_{\mathbf{p}_i}), i = 1, \dots, N^{co}\}, \quad (3)$$

where $R_{\mathbf{p}_i} = C_l(x_i, y_i)$ is a realization of the experimentally measured lift coefficient for the input flap position $\mathbf{p}_i = (x_i, y_i)$, and N^{co} is the number of input-output pairs in the construction sample. Although the general surrogate framework can handle noisy outputs,¹⁷ the noise contribution is neglected in the work presented in this paper. Information from prior studies, outside sources, or asymptotic behavior can also be incorporated into the approximation process. It is important to note that the surrogate framework makes no assumptions in regard to the approximation technique and will accept, and assess, any approximation $\mathcal{A}(\mathcal{X}^{co})$. Also, no restriction is placed on either N^{co} or the distribution of the construction sample.

To proceed with the description of the surrogate validation, we first introduce the importance function $\rho(\mathbf{p})$. The importance function serves as a probability density function for the selection of the validation points:

$$\int_{\Omega} \rho(\mathbf{p}) \mathbf{d}\mathbf{p} = 1. \quad (4)$$

The importance function also leads to the notion of a ρ -measure associated with $\rho(\mathbf{p})$: for any subdomain $\mathcal{D} \subset \Omega$,

$$\mu_\rho(\mathcal{D}) = \int_{\mathcal{D}} \rho(\mathbf{p}) \mathbf{d}\mathbf{p} < 1. \quad (5)$$

The ρ -measure of \mathcal{D} is simply the weighted relative M -volume of \mathcal{D} .

With the importance function $\rho(\mathbf{p})$ defined, we form the validation sample set

$$\mathcal{X}^{va} = \{(\mathbf{P}_i, R_{\mathbf{P}_i}), i = 1, \dots, N^{va}\}, \mathbf{P}_i \sim \rho(\mathbf{p}), \quad (6)$$

where the input flap positions \mathbf{P}_i for the validation sample set are drawn randomly according to the probability density function $\rho(\mathbf{p})$. In Equation (6), the \sim should be read as “is drawn according to the probability density function.” The validation sample size N^{va} is given by

$$N^{va} = \frac{\ln \varepsilon_2}{\ln(1 - \varepsilon_1)}, \quad (7)$$

and ε_1 and ε_2 are the two uncertainty parameters described below. The model prediction error U is computed from the validation sample set \mathcal{X}^{va} as

$$U = \max_{\mathbf{P}_i \in \mathcal{X}^{va}} \frac{|R_{\mathbf{P}_i} - \tilde{S}(\mathbf{P}_i)|}{\hat{g}(\mathbf{P}_i)}, \quad (8)$$

where \mathcal{X}^{va} denotes the input points of the validation sample set and $\hat{g}(\mathbf{p})$ is a strictly positive, error-scaling function described in more detail later.

The result of the construction/validation process is a probabilistic statement that describes the global quality of the surrogate $\tilde{S}(\mathbf{p})$. The validation statement can be compactly written as

$$Pr\{\mu_\rho(\Upsilon) < \varepsilon_1\} \geq 1 - \varepsilon_2, \quad (9)$$

where $Pr\{event\}$ is the “probability of *event*” and $\Upsilon \subset \Omega$ is the uncharacterized region defined as

$$\Upsilon = \left\{ \mathbf{p} \in \Omega \mid |\mathcal{S}(\mathbf{p}) - \tilde{S}(\mathbf{p})| > U \hat{g}(\mathbf{p}) \right\}. \quad (10)$$

The ρ -measure of the uncharacterized region is bounded by ε_1 , and the significance level of the non-parametric statistical bound is ε_2 . This result can be readily proved¹² with order statistics.¹⁸

For the simple case of $\hat{g}(\mathbf{p}) = 1$, Equation (9) states that, with probability greater than or equal to $1 - \varepsilon_2$, the surrogate error is bounded by U over a region of Ω of ρ -measure greater than $1 - \varepsilon_1$. Although this statement is suggestive, it gives neither an indication as to the location of Υ nor the magnitude of the surrogate error in Υ .

Surrogate-Based Optimization

For the optimization problem, we assume that we are given Q target drag coefficient values λ^q , $q \in \mathcal{Q} = \{1, \dots, Q\}$. The goal of the optimization is to find the surrogate-predicted flap positions that minimize the objective function,

$$\mathbf{p}^q = \tilde{\mathbf{p}}^*(\lambda^q) = \arg \min_{\mathbf{p} \in \Omega} |\tilde{S}(\mathbf{p}) - \lambda^q|, \quad \forall q \in \mathcal{Q}. \quad (11)$$

The Q targets could represent different target lift coefficients during the flap deployment schedule, or reflect the goals at different flight conditions (e.g., take-off and landing).

A posteriori Error Analysis

To present the predictability results, we must first introduce the notion of a prediction neighborhood. We begin with a pseudometric $\Delta(\mathbf{a}, \mathbf{b})$ defined for all $(\mathbf{a}, \mathbf{b}) \in \Omega \times \Omega$, which determines a “distance” between two input points \mathbf{a} and \mathbf{b} . Then for any subdomain $\mathcal{D} \subset \Omega$ we define the radius of \mathcal{D} about a point \mathbf{p} as $r_{\mathcal{D}}(\mathbf{p}) = \max_{\mathbf{p}' \in \mathcal{D}} \Delta(\mathbf{p}, \mathbf{p}')$. The prediction neighborhood located at point \mathbf{p} with a ρ -measure of z , $\mathcal{P}(\mathbf{p}, z)$, is that (or a) region $\mathcal{D} \subset \Omega$ of ρ -measure z that minimizes $r_{\mathcal{D}}(\mathbf{p})$. We assume that \mathbf{p} lies inside $\mathcal{P}(\mathbf{p}, z)$ and that $\mathcal{P}(\mathbf{p}, z_1) \subset \mathcal{P}(\mathbf{p}, z_2)$ for $z_1 < z_2$. We can then show that, with probability greater than $1 - \varepsilon_2$, for all $q \in \mathcal{Q}$, regions $\Gamma^q \subset \mathcal{P}(\mathbf{p}^q, \varepsilon_1)$ of nonzero measure exist such that for all $\mathbf{p}' \in \Gamma^q$,

$$|\mathcal{S}(\mathbf{p}') - \tilde{S}(\mathbf{p}^q)| \leq e(\mathbf{p}^q). \quad (12)$$

It now remains to bound $e(\mathbf{p}^q)$ and make precise the extent of Γ^q .

Several bounds are possible on $e(\mathbf{p}^q)$, which we denote the predictability gap. If we wish to bound the predictability of each design individually, we find that

$$e(\mathbf{p}^q) \leq \mathcal{E}(\mathbf{p}^q, \varepsilon_1), \quad \forall q \in \mathcal{Q}, \quad (13)$$

where, for $\mathbf{p} \in \Omega$ and $0 < z < 1$,

$$\mathcal{E}(\mathbf{p}, z) = U g(\mathbf{p}, z) + \delta(\mathbf{p}, z), \quad (14)$$

and

$$g(\mathbf{p}, z) = \max_{\mathbf{p}' \in \mathcal{P}(\mathbf{p}, z)} \hat{g}(\mathbf{p}'), \quad (15)$$

$$\delta(\mathbf{p}, z) = \max_{\mathbf{p}' \in \mathcal{P}(\mathbf{p}, z)} |\tilde{S}(\mathbf{p}') - \tilde{S}(\mathbf{p})|, \quad (16)$$

and U is the model prediction error from the validation step, Equation (8).

In addition to the joint estimates to the bound on $e(\mathbf{p}^q)$, we can also bound the average error over the Q target designs. In particular, if we assume that the $\mathcal{P}(\mathbf{p}^q, \varepsilon_1)$ are mutually disjoint, it can be shown that

$$\frac{1}{Q} \sum_{q=1}^Q e(\mathbf{p}^q) \leq \max_{\underline{\beta} \in \mathcal{C}_Q} \left[\frac{1}{Q} \sum_{q=1}^Q \mathcal{E}(\mathbf{p}^q, \beta_q \varepsilon_1) \right], \quad (17)$$

where $\underline{\beta} = \{\beta_1, \dots, \beta_Q\}$, and

$$\mathcal{C}_L = \{ \underline{\beta}' \in \mathbb{R}^L \mid 0 \leq \beta'_l \leq 1, l = 1, \dots, L; \sum_{l=1}^L \beta'_l = 1 \}, \quad (18)$$

is the set of convex L -tuples. The “nonparametric average” is relevant to multiple-target designs and represents the average, as opposed to the worst-case, estimate of the predictability. Also, it is important to note that this predictability bound is calculated entirely in terms of the inexpensive surrogate, $\tilde{\mathcal{S}}(\mathbf{p})$.

Finally, for a successful validation (i.e., $\mu_\rho(\Upsilon) < \varepsilon_1$), we can bound the expectation of the size of Γ^q with respect to the validation sample joint probability density. The resulting bound is, $\forall q \in \mathcal{Q}$,

$$E\left(\frac{\mu_\rho(\Gamma^q)}{\varepsilon_1} \mid \mu_\rho(\Upsilon) < \varepsilon_1\right) \leq 1 + \frac{1}{\ln \varepsilon_2} + \frac{\varepsilon_2}{(1 - \varepsilon_2)}. \quad (19)$$

The expression in Equation (19) bounds the average ρ -measure of the region Γ^q , with respect to ε_1 , for many validations.

Several advantages to bounding the errors only to within a finite uncertainty exist.¹⁹ First, we achieve a sense of stability in that the estimates apply not only to a single point, but to regions Γ^q of nonzero measure, assuring that many input points \mathbf{p}^q exist that satisfy the error estimates. Second, for the multiple-target case the estimates become sharper because there is only a single uncharacterized volume of measure ε_1 . Equation (17) is the upper bound for the distribution of the single ε_1 -sized uncharacterized region among the Q designs. This analysis results in a bound on the average error which is less than the average of the individual predictability gap bounds $\mathcal{E}(\mathbf{p}^q, \varepsilon_1)$. Finally, because our predictability analysis is not premised on any particular set of points, the designer has flexibility in the choice of the metric $\Delta(\mathbf{a}, \mathbf{b})$ (discussed further in the next section).

As mentioned in the introduction, the primary drawback to the surrogate approach is the difficult construction and validation of the surrogate in high dimensional input spaces. We can easily illustrate this point if we consider the uniform importance function $\rho(\mathbf{p})$ and a neighborhood of ρ -measure ε_1 in the input domain $\Omega = [0, 1]^M$. The neighborhood will span at least $\varepsilon_1^{1/M}$ in one of the input directions which rapidly approaches one as $M \rightarrow \infty$. The loss of localization as $M \rightarrow \infty$ produces a corresponding loss in predictability through $\delta(\mathbf{p}, \varepsilon_1)$ in Equation (14). In certain instances, the surrogate approach can be effectively applied to problems with high dimensional input spaces. This includes cases in which the inputs are highly correlated (e.g., for shape optimization where highly oscillatory geometries are not likely optimizers²⁰) or specialized formulations apply (e.g., Pareto formulations²¹). In general however, the surrogate approach is restricted to a moderate number of design variables.

Summary of Surrogate Inputs

To summarize the surrogate framework description, and to highlight the flexibility of the environment, we note that four inputs to the process are determined by the user. These are listed below:

- i. An importance function $\rho(\mathbf{p}) : \Omega \rightarrow \mathbb{R}_+$.
- ii. An error-scaling function $\hat{g}(\mathbf{p}) : \Omega \rightarrow \mathbb{R}_+$.
- iii. Two uncertainty parameters, ε_1 and ε_2 , that satisfy $0 < \varepsilon_1, \varepsilon_2 < 1$.
- iv. A pseudometric $\Delta(\mathbf{a}, \mathbf{b})$.

Each input provides the designer with flexibility, and allows the designer’s experience to impact and improve the final surrogate-predicted designs. Although poor choices for the inputs do not influence the validity of the surrogate results, they greatly reduce the sharpness of the results. A short description and explanation of each input follows.

The importance function $\rho(\mathbf{p})$ reflects the designers prejudices in regard to the regions of Ω that are more likely to contain optimizers. In this context, $\rho(\mathbf{p})$ is essentially a “prior” on $\tilde{\mathbf{p}}^*$. To serve this purpose, $\rho(\mathbf{p})$ is used as the probability density function in the random selection of validation points in Equation (6). A judicious choice of $\rho(\mathbf{p})$ (one that is large in the regions of the final designs and small elsewhere) can significantly increase the sharpness of the *a posteriori* error bounds. The increased sharpness is a consequence of much better physical localization (in terms of input variable extent) of the prediction neighborhood $\mathcal{P}(\tilde{\mathbf{p}}^*, \varepsilon_1)$, which in turn reduces the surrogate sensitivity contribution $\delta(\tilde{\mathbf{p}}^*, \varepsilon_1)$ to the error bound in Equation (14).

The error-scaling function $\hat{g}(\mathbf{p})$ can be used by the designer to reduce the impact of localized surrogate errors on the error bounds of the final design. Because the model prediction error U in Equation (8) is global, a large value of $\hat{g}(\mathbf{p})$ in regions for which the approximation is poor will result in a reduced value of the first term on the right-hand side of Equation (14), provided that the final design does not lie in a region where $\hat{g}(\mathbf{p})$ is large.

The uncertainty parameters ε_1 and ε_2 are related to the number of validation points through Equation (7). This formula allows the precise budgeting of resources and ensures that useful solutions can be obtained. In effect, Equations (7)–(10) describe what is known in a continuous sense about a function based on discrete sampling. Analysis of Equation (7) shows that, asymptotically for small ε_1 and ε_2 , N^{va} increases linearly as ε_1 decreases and only logarithmically as ε_2 decreases. This relationship suggests that

although we can easily (in terms of validation sample size) increase our confidence in the results (smaller ε_2), refining the localization of our results (through smaller ε_1) is much more difficult. The localization has a direct impact on the final error analysis through $\delta(\mathbf{p}, \varepsilon_1)$ in Equation (14). The relative difficulty in further refining the localization illustrates the need to intelligently select $\rho(\mathbf{p})$ and where appropriate, $\Delta(\mathbf{a}, \mathbf{b})$, both of which can have similar effects on the localization error.

The final input to the surrogate approach is the pseudometric $\Delta(\mathbf{a}, \mathbf{b})$. Because $\Delta(\mathbf{a}, \mathbf{b})$ can be chosen post-validation, various metrics can be examined, and the most appropriate selected. One possible trade-off is between design localization (in terms of input variable extent) and predictability in terms of $\delta(\mathbf{p}, \varepsilon_1)$ in Equation (14). An example of the extreme of this trade-off is the sensitivity minimizing metric

$$\Delta(\mathbf{a}, \mathbf{b}) = |\tilde{\mathcal{S}}(\mathbf{a}) - \tilde{\mathcal{S}}(\mathbf{b})| \quad (20)$$

used for the single-point design study of the results section. This metric gives the lowest possible $\delta(\mathbf{p}, \varepsilon_1)$.

Improved Algorithms

Several, more sophisticated surrogate algorithms have been developed^{8,12,17,19–23} but are not described here. First, a surrogate formulation for noisy outputs has been developed.¹⁷ This formulation is clearly appropriate in an experimental setting but is not addressed here. Second, the multiple output case can be efficiently handled, and the formulation can be applied to model selection. Third, elemental decompositions of Ω are possible that yield local errors and allow for rigorous construction/cross-validation schemes.²⁴ Fourth, sequential and adaptive techniques have been developed that allow the incremental deployment of resources to achieve target surrogate accuracies and that more tightly couple the construction and validation phases of the baseline algorithm. Finally, nested validation, in which a hierarchy of models exists (e.g., an extremely expensive “truth” model \rightarrow a high-fidelity model \rightarrow low-fidelity model), has been addressed as well.

Results

To demonstrate the surrogate framework, we have applied it to the experimental design of multielement airfoils; specifically, we are interested in determining the optimal location for the trailing edge flap, based on the lift coefficient C_l in low-speed, high-lift flight regimes. The $M = 2$ design inputs to the problem $\mathbf{p} = (x, y)$ are the x and y positions of the flap,

measured from the leading edge of the main airfoil element and normalized by the main element chord $c_{main} = 14.95$ in. The output of interest is C_l . In addition, several other configuration and flow condition parameters are fixed for the study. These parameters are listed in Table 1 and are the Reynolds number Re , the airfoil angle of attack α , the flap and slat deflection angles δ_{flap} and δ_{slat} , respectively, and the gap and overhang of the slat (expressed as a percentage of the nested chord $c = 18.0$ in.).

In this section, we first describe the method used for the surrogate construction and report the validation results. Second, we consider the single-point design problem of output maximization. Third, we pursue a multiple-target design study which demonstrates the increased sharpness of the nonparametric average error results. Finally, we report the results of on-line optimization studies and compare these results with the off-line, surrogate results.

Surrogate Construction/Validation

The construction sample set \mathcal{X}^{co} consists of 119 input-output pairs that are uniformly spaced on a 17×7 grid. The (x, y) flap positions for the construction sample are plotted as circles in Figure 4. The input domain is divided into three subdomains, $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$, based on the flow conditions over the flap. In the first subdomain Ω_1 , the flow over the flap is attached, with the exception of the extreme aft positions in which some trailing-edge separation may be present (and desirable). In this region, a radial basis function²⁵ serves as the approximation method, which yields the surrogate $\tilde{\mathcal{S}}_1(\mathbf{p})$. In Ω_3 , the flow over the flap is fully separated, and a second radial basis function fit serves as the surrogate $\tilde{\mathcal{S}}_3(\mathbf{p})$. In Ω_2 , the resolution of the construction points is not sufficient to determine the precise location of the separation line. In this region, a simple linear triangulation between $\tilde{\mathcal{S}}_1(\mathbf{p})$ and $\tilde{\mathcal{S}}_3(\mathbf{p})$ is used as the surrogate, $\tilde{\mathcal{S}}_2(\mathbf{p})$. The error function, $\hat{g}(\mathbf{p})$, is set to unity in Ω_1 and Ω_3 , and $\hat{g}(\mathbf{p}) = 50$ in Ω_2 , reflecting our uncertainty in regard to the location of the

Re	1,000,000
α	14°
δ_{flap}	30°
δ_{slat}	−30°
gap_{slat}	2.17%
$overhang_{slat}$	−1.46%

Table 1: Fixed design study parameters.

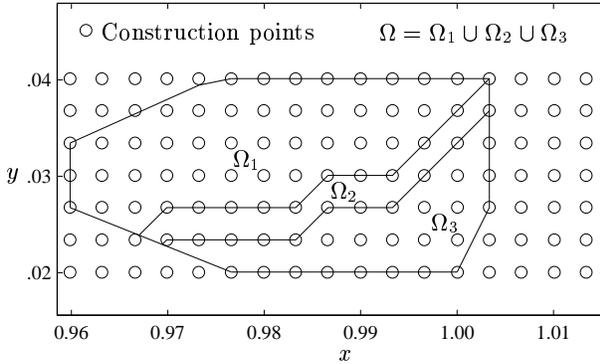


Figure 4: Surrogate construction points and the input (“design”) domain.

separation line and, hence, our lack of confidence in the quality of the surrogate in this region of the input space. A three-dimensional surface plot of the surrogate is shown in Figure 5.

To validate the lift coefficient surrogate, we must select a set of random input points in Ω and run the experiment at each of these points to form the validation sample set \mathcal{X}^{va} . The input points are confined to the design space Ω described in the previous paragraph and shown in Figure 4. Because the construction data were obtained simultaneously with the validation data, we had no expectation in regard to those regions of the input space that would be of most interest; thus, we used a uniform probability density function $\rho(\mathbf{p})$ for the selection of the validation points. We budgeted $N^{va} = 45$ points for validation and, using the relationship in Equation (7), set $\varepsilon_1 = 0.03$ and $\varepsilon_2 = 0.25$. If we had known the form of the surrogate prior to taking the validation data, we could have restricted the design space to a more feasible region and perhaps chosen an importance function $\rho(\mathbf{p})$ that would have concentrated validation points close to potential designs. The scaled model prediction error computed according to Equation (8) is $U = .0482$. Note that the maximum un-scaled error does in fact occur in Ω_2 as we presupposed and has a value of 0.4824. If we had chosen $\hat{g}(\mathbf{p}) = 1$ everywhere (instead of as described above), our model prediction error would have been approximately one order of magnitude larger, and would surely have overwhelmed the results.

The surrogate just described and the related validation results serve for all of the designs discussed in the remainder of this paper. One primary advantage to using the surrogate approach is the fact that no additional experimental data are required to bound the errors of future designs that are pursued with the surrogate. This characteristic, combined with

the negligible computational time required for each surrogate evaluation, yields a highly flexible design environment that does not sacrifice predictability.

Single-Point Design, Surrogate Maximization

For the first study, we pursue a single-point design that maximizes the surrogate output. We set λ sufficiently large in Equation (2) and minimize the resulting function. To accomplish the optimization, we use the unconstrained quasi-Newton optimizer that is included in the optimization toolbox of Matlab²⁶. The resulting surrogate-based optimizer is located at $\tilde{\mathbf{p}}^* = (x^*, y^*) = (.997, .036)$, and the surrogate-predicted lift coefficient value at this point is $\tilde{\mathcal{S}}(\tilde{\mathbf{p}}^*) = 3.388$. The optimizer was started with an initial guess at $\mathbf{p}_0 = (.987, .033)$ and required 44 surrogate evaluations to arrive at $\tilde{\mathbf{p}}^*$. Because the surrogate is inexpensive to evaluate (and because we are working with only two inputs and can visualize the results graphically), we can verify that we do achieve a surrogate-predicted global maximum. This verification would be more difficult in a purely on-line optimization setting if we did not begin the optimizer at multiple starting points \mathbf{p}_0 until we had sufficient confidence that a global maximum had been obtained.

Finally, we choose the sensitivity minimizing metric $\Delta(\mathbf{a}, \mathbf{b}) \equiv |\tilde{\mathcal{S}}(\mathbf{a}) - \tilde{\mathcal{S}}(\mathbf{b})|$ in Equation (20) and perform the *a posteriori* error analysis for a single-point design. We construct the prediction neighborhood $\mathcal{P}(\tilde{\mathbf{p}}^*, \varepsilon_1)$ around $\tilde{\mathbf{p}}^*$ and find the surrogate sensitivity parameter $\delta = .0328$. The optimal point $\tilde{\mathbf{p}}^*$ and the associated prediction neighborhood $\mathcal{P}(\tilde{\mathbf{p}}^*, \varepsilon_1)$ are plotted in Figure 6. The resulting predictability statement reads as follows: with confidence level greater than .75, a region $\Gamma \subset \mathcal{P}(\tilde{\mathbf{p}}^*, \varepsilon_1)$ of nonzero measure exists such that for all $\mathbf{p}' \in \Gamma$

$$|\mathcal{S}(\mathbf{p}') - \tilde{\mathcal{S}}(\tilde{\mathbf{p}}^*)| \leq e(\tilde{\mathbf{p}}^*), \quad (21)$$

where

$$e(\tilde{\mathbf{p}}^*) \leq Ug(\tilde{\mathbf{p}}^*, \varepsilon_1) + \delta = .0810. \quad (22)$$

We see that the predictability is relatively good with respect to the surrogate-predicted maximum lift coefficient, but quite poor with respect to the range of lift coefficients of interest (i.e., corresponding to flap positions in Ω_1).

Multiple-Target Designs

For the second design study, we pursue a multiple-target design. The motivation for such a study might be an interest in examining the lift coefficient at more than one point of the deployment of the flap.

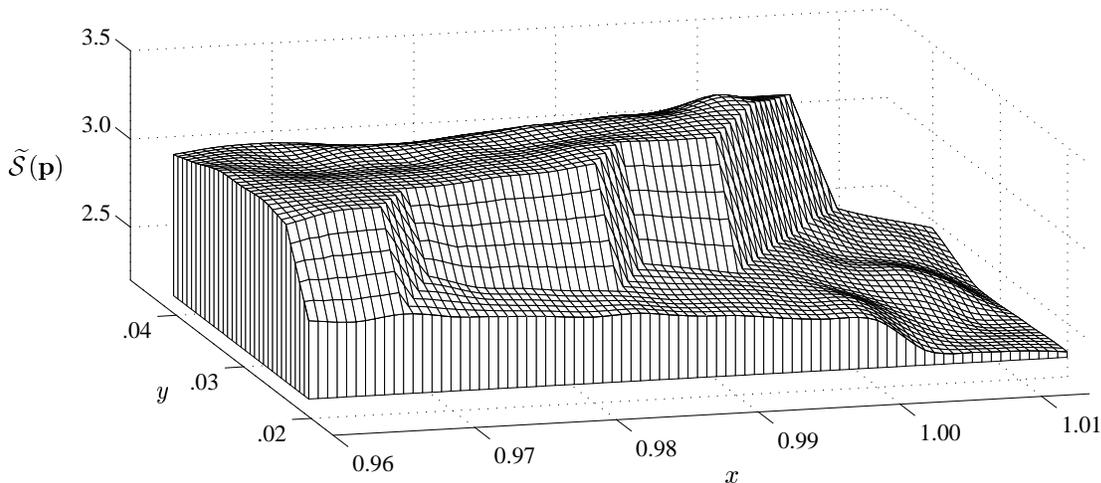


Figure 5: Three-dimensional mesh plot of the lift coefficient surrogate $\tilde{S}(\mathbf{p})$.

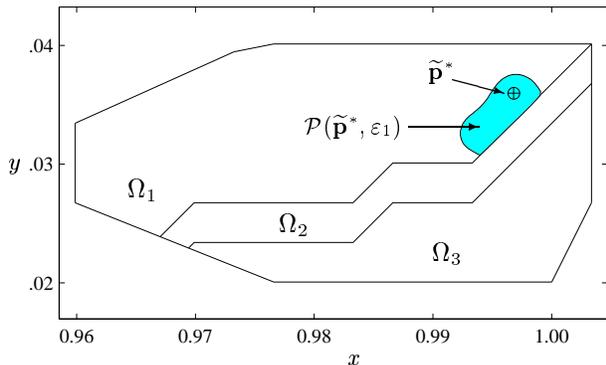


Figure 6: The surrogate-predicted optimizer, $\tilde{\mathbf{p}}^*$, and the associated prediction neighborhood, $\mathcal{P}(\tilde{\mathbf{p}}^*, \varepsilon_1)$.

Specifically, we want to obtain two target lift coefficients: $\lambda^1 = 3.31$ and $\lambda^2 = 3.25$. Isocontours of the surrogate indicate that a locus of points in Ω exists for each target that exactly satisfies the design goals. We arbitrarily select one point for each design: $\mathbf{p}^1 = (x^{[1]}, y^{[1]}) = (.987, .033)$ and $\mathbf{p}^2 = (x^{[2]}, y^{[2]}) = (.979, .033)$. Around each optimizer, we construct a prediction neighborhood chosen from the family of ellipses that have area equal to ε_1 , are centered at \mathbf{p}^q , and are oriented such that they minimize surrogate sensitivity $\delta(\mathbf{p}^q, \varepsilon_1)$. The optimizers and associated prediction neighborhoods are plotted in Figure 7.

For each of the designs ($q = 1, 2$), we can state with confidence level greater than .75 that a region $\Gamma^q \subset \mathcal{P}(\mathbf{p}^q, \varepsilon_1)$ of nonzero measure exists such that for all $\mathbf{p}' \in \Gamma^q$

$$|\tilde{S}(\mathbf{p}^q) - S(\mathbf{p}')| < e(\mathbf{p}^q), \quad (23)$$

where

$$e(\mathbf{p}^1) = U + \delta(\mathbf{p}^1, \varepsilon_1) = .0482 + .0198 = .0680, \quad (24)$$

and

$$e(\mathbf{p}^2) = U + \delta(\mathbf{p}^2, \varepsilon_1) = .0482 + .0201 = .0683. \quad (25)$$

The above bounds jointly hold on each design. We obtain a slightly sharper bound on the average error of the two designs:

$$\frac{1}{2}[e(\mathbf{p}^1) + e(\mathbf{p}^2)] \leq U + .0149 = .0631. \quad (26)$$

The increased sharpness results from an analysis of the worst-case distribution of the uncharacterized region between the two prediction neighborhoods. Because of the low sensitivity of the surrogate in each of the prediction neighborhoods relative to model prediction error U , the improvement is slight.

Comparison with Direct Insertion

To date, cases at identical flow conditions have not been examined with both on-line (the method of steepest ascent) and off-line (the surrogate approach) optimization methods. However, rough comparisons of the resource requirements are of (guarded) use.

The on-line results have been reported in an earlier paper by Landman and Britcher.⁴ In that effort, they found the optimizer to be very robust (successful in 6 out of 6 attempts) and insensitive to the initial guess. For each case, they started the optimizer at in initial flap position with a low C_l value and obtained a final value within approximately 0.7 percent of the maximum C_l value in approximately 20 optimizer steps, requiring approximately 60 experimental

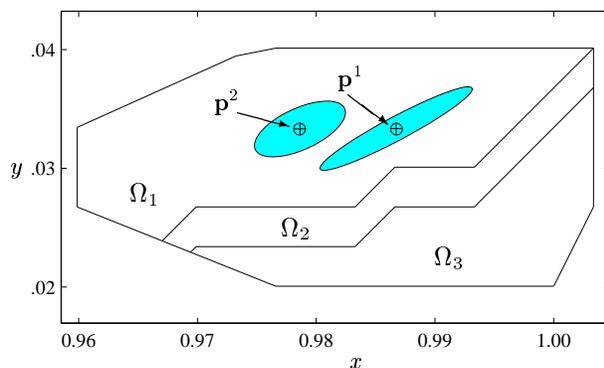


Figure 7: The surrogate-predicted optimizers and the associated prediction neighborhoods (shaded).

data points (3 points per step). With the surrogate method, we required 119 points to construct the surrogate and an additional 45 for the validation, for a total of 164 experimental data points. For the maximization problem, the *a posteriori* error bound was 2.4 percent of the maximum surrogate value.

While the surrogate approach seems to compare unfavorably to the on-line method, several subtleties lie in its favor. First, for designs chosen with the validated surrogate in the future (e.g., the multiple-target design examined in this paper), similar error bounds still apply and do not require additional experimental data. In contrast, the on-line approach would require additional experimental results. Second, a total of 60 evaluations to obtain an optimal point with the on-line method can be deceptive; to be assured that the result is indeed optimal, additional information is required. The additional information for the study cited was in the form of contour plots of a matrix of data. If visualization is not possible, a number of optimizer restarts would be required to be assured of an optimal. Third, in cases for which the objective function is less forgiving, restarts of the on-line optimizer would be unavoidable, which would further increase the required experimental data to a level surpassing that of the surrogate approach. Finally, the obvious difficulty in pursuing on-line optimization is related to the ultimate application; if the intent is to incorporate the data as a portion of a larger optimization study, no alternative is available other than to store the experimental data for later use and extract with some form of an approximation. If one is restricted to a purely experimental setting, then the ability to quickly, and automatically, find optimal operating points with the on-line optimizer is highly advantageous.

Acknowledgments

This work has been partially supported by two ASEE Summer Faculty Research Fellowship Awards (D.L.), by NASA Langley Research Center (LaRC) under task order NAS1-19858 (D.L.), and currently under grant NAG1-1750, with technical monitor Dr. John C. Lin (D.L.). Additionally, this work has been supported by the Defense Advanced Research Projects Agency under grant N00014-91-J-1889 (A.T.P.), by the Air Force Office of Scientific Research under grant F49620-94-1-0121 (A.T.P.), and by NASA under grant NAG 1-1613 (A.T.P.). Finally, we would like to acknowledge the NASA LaRC Graduate Training Program (J.C.O.).

References

- ¹ Landman, D. and Britcher, C. P., “Advanced Experimental Methods for Multi-Element Airfoils,” AIAA Paper 95-1784, 1995.
- ² Fox, R. L., *Optimization Methods for Engineering Design*, Addison Wesley, 1971.
- ³ Box, G. E. P. and Draper, N., *Empirical Model Building and Response Surfaces*, John Wiley & Sons, New York, 1987.
- ⁴ Landman, D. and Britcher, C. P., “Experimental Optimization Methods for Multi-Element Airfoils,” AIAA Paper 96-2264, 1996.
- ⁵ McKay, M. D., Beckman, R. J., and Conover, W. J., “A Comparison of Three Methods For Selecting Values of Input Variables In the Analysis of Output from a Computer Code,” *Technometrics*, 21, 1979, pp. 239-245.
- ⁶ Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P., “Design and Analysis of Computer Experiments,” *Statistical Science*, Vol. 4, 1989, pp. 409-435.
- ⁷ Barthelemy, J.-F. M. and Haftka, R. T., “Approximation Concepts For Optimum Structural Design — A Review,” *Structural Optimization*, Vol. 5, 1993, pp. 129-144.
- ⁸ Yesilyurt, S., and Patera, A. T., “Surrogates For Numerical Simulations; Optimization of Eddy-Promoter Heat Exchangers,” *Comp. Methods Appl. Mech. Engr.*, Vol. 121, 1995, pp. 231-257.
- ⁹ *LabView for Windows User Manual*, National Instruments Corporation, 1993.
- ¹⁰ Nakayama, A. Kreplin, H.-P., and Morgan, H. L., “Experimental Investigation of Flowfield About a Multielement Airfoil,” *AIAA Journal*, Vol. 28, No. 1, 1990, pp. 14-21.

- ¹¹ Seber, G. A. F. and Wild, C. J., *Nonlinear Regression*, John Wiley & Sons, New York, 1989.
- ¹² Yeşilyurt, S., “Construction and Validation of Computer-Simulation Surrogates For Engineering Design and Optimization,” Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, 1995.
- ¹³ Bohlin, T., *Interactive System Identification: Prospects and Pitfalls*, Springer-Verlag, Berlin, 1991.
- ¹⁴ Valiant, L. G., “A Theory For the Learnable,” *CACM*, Vol. 27, 1984, pp. 1134-1142.
- ¹⁵ Gallant, S. I., “A connectionist learning algorithm with provable generalization and scaling bounds,” *Neural Networks*, **3**, 1990, pp. 191-201.
- ¹⁶ Traub, J. F., Wasilkowski, G. W., and Woźniakowski, H., *Information-Based Complexity*, Academic Press, San Diego, 1988.
- ¹⁷ Yeşilyurt, S., Ghaddar, C., Cruz, M., and Patera, A. T., “Bayesian-Validated Surrogates For Noisy Computer Simulations; Application To Random Media,” *SIAM J. Sci. Comput.*, to appear, 1996.
- ¹⁸ David, H. A., *Order Statistics*, 2nd ed., John Wiley & Sons, New York, 1981.
- ¹⁹ Otto, J. C., Paraschivoiu, M., Yeşilyurt, S., Patera, A. T., “Bayesian-Validated Computer-Simulation Surrogates For Optimization and Design: Error Estimates and Applications,” *Proceedings, IMACS-COST conference at EPFL, J. Appl. Numer. Math*, 1996.
- ²⁰ Otto, J. C., Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, in progress, 1996.
- ²¹ Kambourides, Miltos E., Masters Thesis, Massachusetts Institute of Technology, Cambridge, MA, in progress, 1996.
- ²² Otto, J. C., Paraschivoiu, M., Yeşilyurt, S., Patera, A. T., “Bayesian-Validated Computer-Simulation Surrogates For Optimization and Design,” *Proceedings, ICASE Workshop on Multidisciplinary Design Optimization*, Hampton, VA, SIAM, 1995.
- ²³ Paraschivoiu, M., Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, in progress, 1996.
- ²⁴ Stone, M., “Cross-Validatory Choice and Assessment of Statistical Predictions,” *J. Roy. Stat. Soc. Ser. B*, Vol. 36, 1974, pp. 111-147.
- ²⁵ Dyn, N., Levin, D., and Rippa, S., “Numerical Procedures for Surface Fitting of Scattered Data by Radial Functions,” *SIAM J. Sci. Comput.*, Vol. 7, No. 2, 1986.
- ²⁶ *MATLAB Reference Guide*, The MathWorks Inc., Natick, MA, 1992.

**A SURROGATE APPROACH TO THE EXPERIMENTAL OPTIMIZATION
OF MULTIELEMENT AIRFOILS**

John C. Otto

Multidisciplinary Optimization Branch
NASA Langley Research Center
Hampton, Virginia
Department of Aeronautics and Astronautics, M.I.T.

Drew Landman

Department of Engineering Technology
Old Dominion University
Norfolk, Virginia

Anthony T. Patera

Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts

July 12, 1996

To appear in the Proceedings of the Sixth AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, September 4-6, 1996, Bellevue, WA. Address all correspondence to: John C. Otto, Mail Stop 159, NASA Langley Research Center, Hampton, VA 23681-0001, U.S.A.