UNIVERSITY OF Southampton
School of Engineering Sciences

# Multi-fidelity global optimization using co-kriging[1]


UTC for Computational Engineering
Alexander Forrester, András Sóbester and Andy Keane: CEDG School of Engineering Sciences



## Abstract

Optimization problems, where the set of conditions are sought which relate to an optimum scenario, are encountered in most branches of science. In this work we consider the class of optimization problems where there are a number of ways of obtaining values of the quantity to be optimized (the "objective function"). Data from fast (but less trustworthy) and slow (but more accurate) calculations or measurements can be correlated to learn more about the objective function. We present an optimization strategy which uses correlated data to build a statistical model of the objective function to predict the location of the global optimum.


## A one variable example

Imagine that our expensive to compute data are calculated by the function $f_e(x)=(6x-2)^2\sin(12x-4)$, $x\in[0,1]$, and a cheaper estimate of this data is given by $f_c(x)=Af_e+B(x-0.5)-C$. We sample the design space extensively using the cheap function at $\mathbf{X}_c=\{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$, but only run the expensive function at four of these points, $\mathbf{X}_e=\{0, 0.4, 0.6, 1\}$.

**Fig. 1** shows the functions $f_e$ and $f_c$ with $A=0.5$, $B=10$, and $C=-5$. A kriging prediction through $y_e$ gives a poor approximation to the deliberately deceptive function, but the co-kriging prediction lies very close to $f_e$, being better than both the standard kriging model and the cheap data. Despite the considerable differences between $f_e$ and $f_c$, a simple relationship has been found between the expensive and cheap data and the estimated error reduces almost to zero at $\mathbf{X}_c$ (see **Fig.2**).

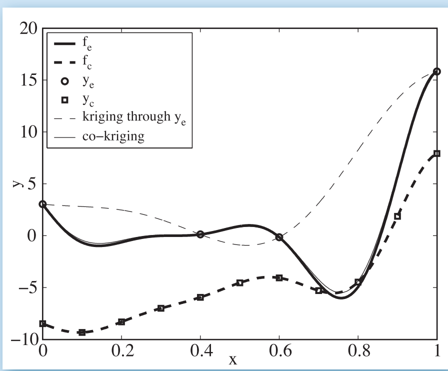


**Fig.1** A one variable co-kriging example. The kriging approximation using four expensive data points ($y_e$) has been significantly improved using extensive sampling from the cheap function ($y_c$).

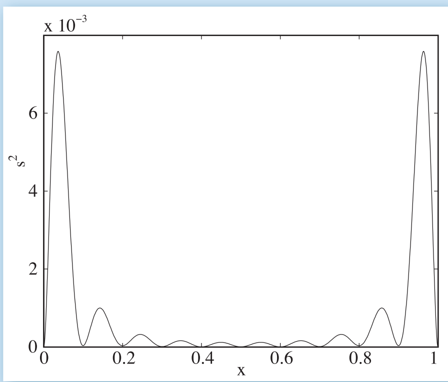


**Fig.2** Estimated error in the co-kriging prediction in figure 1. The simple relationship between the data results in low error estimates at $\mathbf{X}_c$ as well as $\mathbf{X}_e$.

## An aerospace example

The benefits of co-kriging, and in particular regressing co-kriging, are demonstrated through the optimization of a generic transonic civil aircraft wing. Our cheap analysis is in the form of an empirical drag estimation code, *Tadpole*. This returns a drag value based on curve fits to previously analysed wings in ~ 0.6 s. Our 'expensive' code is the linearised potential method *VSaero* with viscous coupling (~2 min per drag evaluation). We aim to minimise drag/dynamic pressure ($D/q$) for a fixed lift (determined by a wing weight calculated by Tadpole).

To allow visualisation of the design landscape, we limit the search to the four variables which have the most impact on drag: area ($S$), aspect ratio AR, sweep $\Lambda$, and inboard taper ratio ($T_{in}$). The Tadpole and VSaero design landscapes are shown using hierarchical axis plots in **Fig.3** and **Fig.4**. Each tile of the plots shows $D=q$ for $S\in[150, 250]\text{m}^2$ and AR $[6, 12]$ for a fixed $\Lambda$ and Tin. $\Lambda$ and Tin vary from tile to tile with the value at the lower left corner of the tile representing the value for the entire tile. The blank portions of **Fig. 4** are where VSaero has failed to return a result for unusual geometries which lead to extreme flow regimes. We start with an initial sampling plan for $\mathbf{X}_c$ of 100 points with a subset of $n_e = 20$ points from which to build the initial co-kriging model. $\mathbf{X}_c$ and $\mathbf{X}_e$ are shown in **Fig. 3**. Despite the apparent sparseness of the $\mathbf{X}_e$ data, a good initial co-kriging prediction of the VSaero landscape is obtained, with a correlation coefficient of 0.96 when compared to the $11^2\times 3^2$ VSaero data-set. We follow an iterative process of updating the co-kriging model with new VSaero and Tadpole data at points which maximize the expectation of improvement and re-tuning hyper-parameters until the optimum is found.

The co-kriging based search consistently outperforms a kriging based search: finding better optima for reduced numbers of VSaero evaluations and with fewer failed simulations. The initial prediction of the kriging model is almost as accurate as the co-kriging model (see correlation and RMSE in **Table 1**), but the greater coverage of data in the co-kriging model leads to better selection of successful update points in promising regions, as seen in **Fig. 4**, which shows the distribution of update evaluations for a typical search. Moreover, the co-kriging updates are concentrated in regions of good designs, while the kriging updates are more widespread because, as there is a sparsity of data, there is high error and therefore high expectations of improvement in many areas (i.e. there is an emphasis on exploration over exploitation).

## Conclusions

Our results demonstrate that correlating analyses at multiple levels of fidelity can enhancethe accuracy of a surrogate model of the highest level of analysis. This correlated model can be used to find optimal solutions more quickly. We have presented a global optimization strategy using a co-kriging based method which allows for varying levels of noise filtering across multi-fidelity analyses and converges towards the global optimum using expected improvement maximization.

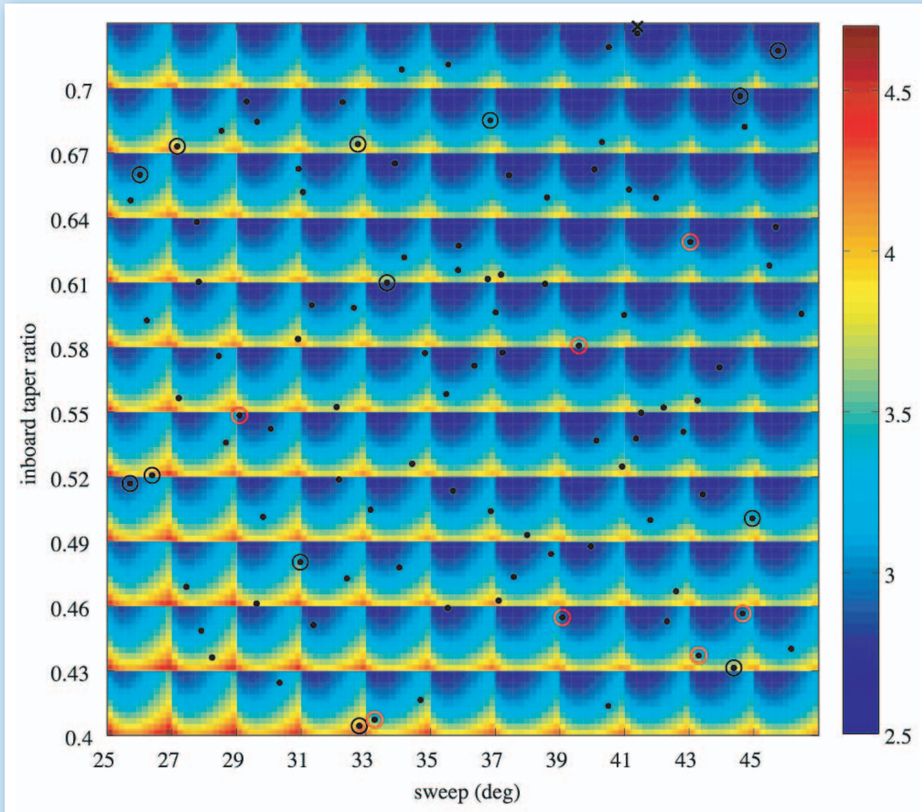


**Fig.3** Tadpole calculated $D/q$ with a typical sampling plan. $\mathbf{X}_c$ (Tadpole evaluations) are shown as dots and are circled at locations of $\mathbf{X}_e$ (VSaero evaluations). Red circles indicate failed VSaero simulations.

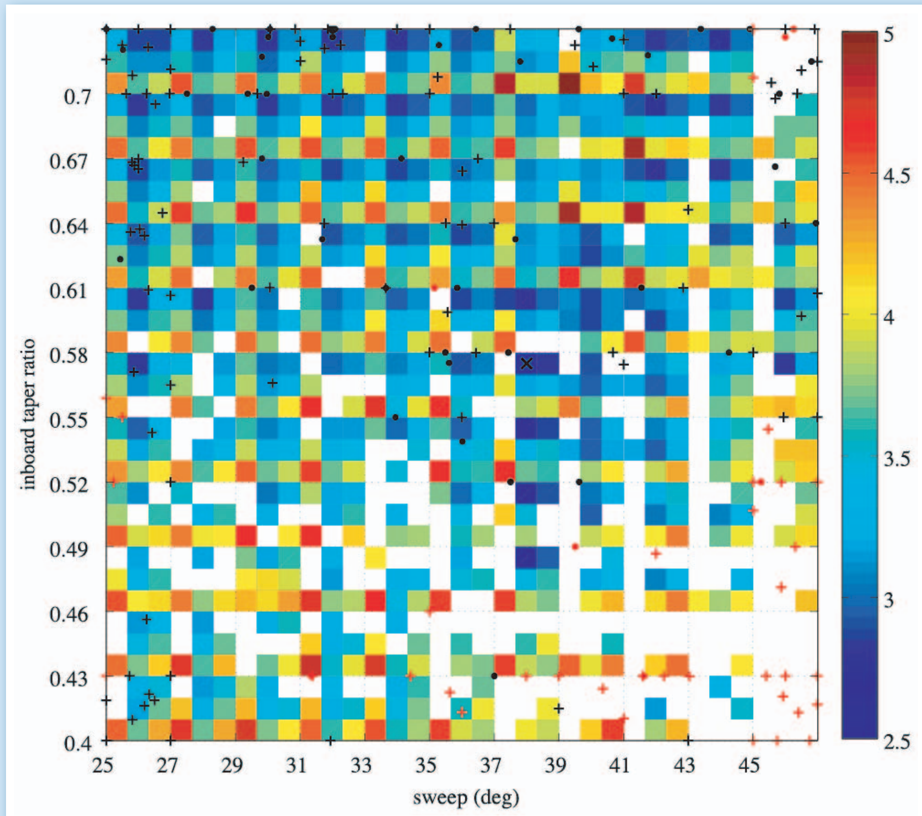


**Fig.4** VSaero calculated $D/q$ with co-kriging (dots) and kriging (crosses) updates from a typical search (the search which yielded values closest to the mean results in **Table 1**).

| | initial model | | best D/q ($m^2$) | function evaluations | | |
|---|---|---|---|---|---|---|
| | $r^2$ | RMSE | VSaero | Tadpole | suc. VSaero | failed VSaero |
| kriging | 0.949 | 0.155 | 2.621 | 0 | 119 | 40 |
| co-kriging | 0.962 | 0.108 | 2.556 | 185 | 79 | 27 |

**Table 1** Performance comparison for the four variable transonic wing problem (averaged over five searches).

[1] This poster is based on the article "Multi-fidelity optimization via surrogate modelling", Alexander I. J. Forrester, András Sóbester and Andy J. Keane, *Proc. R. Soc. A*, 463(2088), 3251-3269, (doi:10.1098/rspa.2007.1900).


**www.soton.ac.uk/ses/research/ced/posters.html** | email: Alexander.Forrester@soton.ac.uk
Computational Engineering & Design Group, University of Southampton, SO17 1BJ, U.K.