Towards Optimization by Matching of Response Surfaces: Finding Windows of Maximal Similarity

Verónica Díaz Pacheco, Mary Acosta and Mauricio Cabrera-Rios
Industrial Engineering Department
University of Puerto Rico - Mayagüez
Mayagüez, PR, USA
veronica.diaz4@upr.edu

Abstract

The ultimate goal of this work is to find a region where the response surface of a function that is not well characterized in terms of optimality resembles one that is well-characterized in such terms to find, at least, a local optimum. The region in the input space where this resemblance occurs, we call a Window of Maximum Similarity (WMS) and is identified by formulating and solving an optimization problem. The method is one of minimization of squared errors and can be used to explore experimental or simulated data. A series of examples are presented to show the method's feasibility and capability for generating a two-dimensional WMS. This method is a viable element that will serve for the future development of Optimization by Similarity.

Keywords
Optimization, Simulation, Metamodels, Response Surface Methodology

1. Introduction
Typical approaches when using modeling require identifying an equation that can best approximate a dataset. For example, in linear regression, data is used to fit the line that minimizes the sum of squared errors between data points-the functions input space- and the fitted line- its expected response. Even so, responsible statistical practice requires analyzing those points, in statistics called 'outliers', where the model seems not to explain the independent variable as well. In this space, data does not behave in a desirable manner, although there will ideally exist a region in the dataset where it does. Our statement is that there can be regions where data, any data at hand, behaves, or "looks like" a function with desirable optimality conditions and that this can be used to explore the simulated data to find at least a local optimum.

When a complex model, like those frequently used in simulation, is approximated by another simplified one for statistical testing, computational efficiency or optimization purposes, we are using a metamodel. It is common in metamodeling that the model's parameters are estimated until the most competitive fit is found (typically) by minimization of an error function. Metamodels are often used for optimization purposes; after a model form is determined, it can be used to experimentally find optimality conditions of a process, as in Response Surface Methodology. In this work, metamodels are used in a different manner but also pointing towards optimization; after generating data with a model, the metamodel's parameters will be fitted to be the one that maximizes the similarity between itself and the data. By superimposing both functions and finding their region of maximum resemblance, we are looking to identify a window the input space where this similarity occurs.

Inverse mappings, when a function's input is a specific desired performance and its output its associated controllable variable settings, was approached in Barton's (2005) work. From the intricacies he mentioned, it was noted that the task of inverse mappings is often reduced to finding one (or more) input parameter combinations for one certain output characteristic. As in the method here proposed, solving inverse problems by the identification of the regions, instead of points, was assessed in the Couckuyt et al.'s (2010) work. The Window of Maximum Similarity (WMS) method differs from the latter in the sense that it was constructed to be applicable to detect zones of interest in different kinds of data and does not use probability density functions, but rather least squares estimation and linear programming
.

As was first done in Rivera-Nazario and Cabrera-Ríos's (2013) work, our Optimization by Similarity method is applied to search for a region where a metamodel fits best. Their study addresses a common problem faced in modeling polymers: the relationship between deformation and viscosity. In contrast, we propose applying the method to any problem, that is, any that requires modeling, abstracting it to the mathematical space of functions. We also consider a two-dimensional input, or 'controllable variable' space, as opposed to only one. Our method entails matching a (simulated) function -one that represents, or rather, generates random data- to another one that has desired optimality properties-a specific form - and find their region of maximum resemblance, through least squares estimation and optimization, where there could exist at least a local optimum. The development of the method is described below, and its applicability is tested on several common global optimization test functions and a function created by our research group; *AOG_1*.
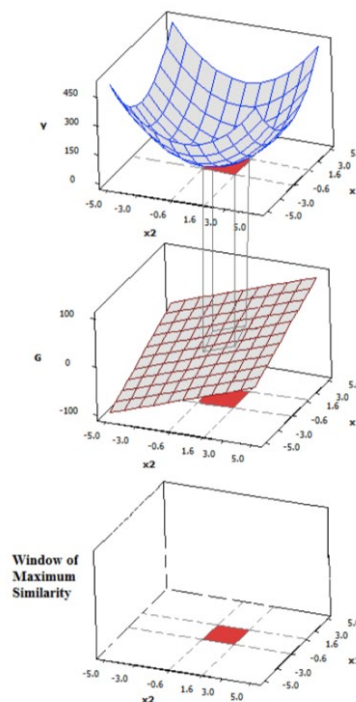

Figure 1: Example for construction of the Optimization by Matching of Responses where function to approximate data is quadratic and function to superimpose is linear; region in red represents Window of Maximal Similarity.

2. Techniques used in this work

2.1 Least Squares

The least squares method is typically used to estimate regression parameters by minimization of the sum of squared errors (SSE). Let the *experimental region*, **R,** be the ith-dimensional hyper-space made up of all possible values that each input variable can take;

$$\boldsymbol{R} : \{ f(x_i) \mid x_i \in [\, x_i^{min}, x_i^{max} \,] \}$$

In this work, the SSE is given by:

$$SSE = \sum_{j=1}^{n} (Y(\boldsymbol{R}) - Z(\boldsymbol{R}))^2$$

Where $Y = f(\boldsymbol{R})$, is the response of the function to approximate that needs to be optimally addressed and $Z = f(\boldsymbol{R})$ is the response of the model or function to superimpose, which has desired and well-established optimality properties, i.e. it is convex and has a global optimum.

2.2 Experimental region discretization

To generate the grid of experimental points used in the proposed method, a discretization size, or step, Δx can be chosen when the input variable initialization values are selected not to be integers. This Δx value can be user-defined, and its use is presented later on in the evaluation of the method using global optimization test functions.

2.3 Multiple starting points

The multiple starting points technique, a heuristic method, is frequently used in order to increase the chance of finding an attractive solution close to the global optimum. When a local optimization method is used, this method is executed many times using different starting points to increase the chance of convergence to a competitive solution (Frontline Systems Inc., 2015a).

3. Proposed method

The aim of this work is to find a region where the response of function that is not well characterized in terms of optimality resembles another one that is well-characterized in such terms, to find at least a local optimum. This well-characterized function, $Z = f(\boldsymbol{R})$, could have a determined static form, or be, ideally, an adjustable metamodel like, for example, a second order polynomial regression with unknown parameters. Once the forms are picked, an experimental region for both functions has to be defined. The region where the maximum similarity occurs is identified by formulating an optimization problem (shown below) which minimizes the functions' SSE. The problem's solution will be the location and size of a window in the experimental region where the maximum similarity between the responses occurs. The optimization problem formulation and an example are illustrated below:

Find x_i^L, x_i^U, to

Minimize $\frac{1}{3} ([log\,(SSE + 1)] - [log(|x_1^U - x_1^L| + 1)] - [log(|x_2^U - x_2^L| + 1)])$ *(1)*

Subject to
$$x_{min}^L \leq x_i^L \leq x_{max}^L \quad *$$
$$x_{min}^U \leq x_i^U \leq x_{max}^U \quad *$$
$$x_i^U - x_i^L \geq \varepsilon \quad ** \quad i = [1,2]$$

*Experimental region and distance between window bound constraints, respectively.

**For the integer variable cases, add the restriction: $x_i^L, x_i^U =$ integer to the model
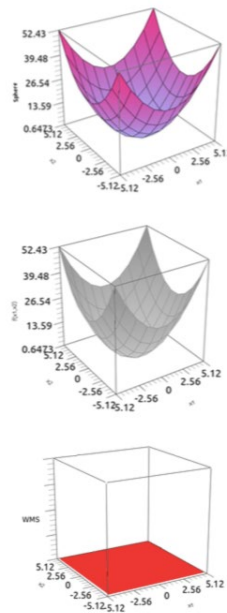
Figure 2: Example of Window of Maximal Similarity found between sphere and second order polynomial regression.

In Figure 1, an example is presented to demonstrate the application of the proposed method, where function to approximate, $Y = f(\boldsymbol{R})$, is the sphere and the function to superimpose, $Z = f(\boldsymbol{R})$, is a second order polynomial regression of the form:

$$Z = f(x_1, x_2) = \beta_0 + \beta_1 x_1^2 + \beta_2 x_2^2.$$

When the function to superimpose has the form of a metamodel, the optimization model has to additionally **find** metamodel parameters $\beta_0, \beta_1, \dots \beta_n$. For the special case where the two functions have the exact same shape, as was the case presented in example 1, the optimization problem has infinitely many solutions since all possible window sizes and locations in the experimental region provide an SSE value of 0. Hence the maximum similarity was found to be all the experimental region, which is precisely the region where both functions are the same.

The logarithm base 10, which is not defined for 0 or negative values, was used in the objective formulation to keep in the same order of magnitude between the SSE and the distance between bounds. Distances between bounds $(x_i^U - x_i^L)$ are present:
1. in the objective function in order to avoid window size dependency on parameters
2. in the constraints because a minimum value must be included for the formulation to be successful and not contain a single point. This suggested value, is $\varepsilon = 1 \times 10^{-6}$. This value is known as the non-archimedean constant, a value commonly used for computational purposes.

Additional constraints include a range to locate the window's upper and lower bounds, respectively, in accordance to the range in which each function varies. In the case that a metamodel is used, an additional constraint must include the range for the metamodel parameter estimates $\beta_0, \beta_1, \dots \beta_n$.

An important quality of the method is its use of computational resources; all the optimization problems included in this work were solved using Excel Solver, a local optimizer included in MS Excel. MS Solver uses the Generalized

Reduced Gradient (GRG) algorithm to solve non-linear optimization problems and the Branch and Bound method to solve mixed-integer and constraint programming problems (Frontline Systems Inc., 2015a and 2015b).

4. Method Evaluation

Two scenarios for evaluation of the method were considered. First, the evaluation of the method using a function to approximate which was designed in our research group, *AOG_1*, is presented. Lastly, a case of an application of the proposed method using unconstrained global optimization test functions is presented. The figures were generated in QtiPlot software (version 0.9.8.9) (http://www.qtiplot.com/).

4.1 AOG_1

It was of our interest to find the maximum similarity between function *AOG_1* and a quadratic function with the form of a bowl because it was reasonable to understand that, potentially, the resulting WMS will match the curve region of the function *AOG_1* with the quadratic function.

    i.   The **function to approximate**, *AOG_1* is a piece-wise function which mostly has the form of a plane except for a given interval in the central experimental region where it looks like a bowl, as shown in figure 3. The ranges x_1 and x_2 were [-5, 5]. *AOG_1* is given by:

$$Y = f(x_1, x_2) = \begin{cases} 5x_1^2 + 5x_2^2 & if \ x_1 \in [-2, 1] \,, x_2 \in [-3, 0 \,] \\ \\ 500 - 5x_1 + \ 5x_2 \ for \ any \ other \ point \end{cases}$$

    ii.   The **function to superimpose** was fixed, for this case, and defined by: $Z = f(x_1, x_2) = \ 5x_1^2 + \ 5x_2^2$. The ranges x_1 and x_2 were also [-5, 5].

    iii.   The **experimental region** is a grid that contains 121 points*. $x_1 \in [-5,5], x_2 \in [-5,5]$,

*$(\, |x_1{}^L - \, x_1{}^U| + 1 \,) \, \times \, (\, |\, x_2{}^L - \, x_2{}^U| + 1 \,)$ = total number of points in experimental region. For this particular case $(\, |-5 - 5| + 1 \,)^2 = 11^2 = 121$

    iv.   The **optimization problem** is given by (1).

    v.   Optimize the model. The global minimum for *AOG_1* is given by solution (0, 0) and its objective value is 0. Essentially, the maximum similarity would be found if the WMS is adjusted within the quadratic region of both functions. The parameters for setting up the Solver that were used for this evaluation include:
- The use of multiple starting points using a population size of 100.
- A constraint precision of 1×10^{-7}.
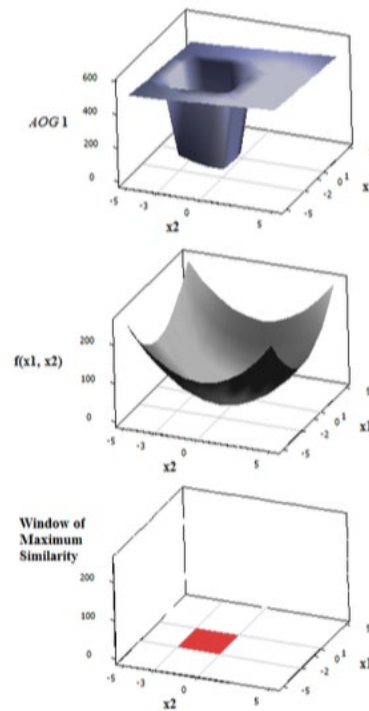- A convergence of 1×10^{-4}.

Figure 3: Function AOG_1 to approximate, quadratic function $Z = f(x_1, x_2) = 5x_1^2 + 5x_2^2$ to superimpose, and Window of Maximum Similarity found for proposed method.

Table 1 reports the best solutions for function *AOG_1*. The type of variables (ToV) is included since evaluations considered the bounds $(x_1^U, x_1^L, x_2^U, x_2^L)$ as both continuous and integer variables. Also, the best solution, the best objective function value (OFV) found, WMS dimensions and location for each case are reported. The function *AOG_1*, the quadratic function $Z = f(x_1, x_2)$ to superimpose, and the WMS of the composite objective function for the integer case are presented in Figure 3. In all cases, the stationary point (0, 0) is *within* the WMS. According to the results, it is possible to conclude that the method demonstrated potential to find regions of similarity between two responses, where optimality can be a pattern of interest.

Table 1: Solutions obtained for function to approximate, AOG_1, and a quadratic function to superimpose.

Type of Variable	Best Solution				Best Objective Value Found				WMS Size
	x_1^L	x_1^U	x_2^L	x_2^U	Objective Function Value	SSE	$x_1^U - x_1^L$	$x_2^U - x_2^L$	
C	-2.17	1.11	-3.73	0.35	-0.45	0	3.28	4.08	3.28 x 4.08
I	-2	1	-3	0	-0.40	0	3	3	3 x 3

© IEOM Society International

4.2 Global optimization test functions

Considered in a two-dimensional input space, the unconstrained global optimization test functions in which the method was tested on included: Sphere, Rosenbrock, Rastrigin, Griewank, Goldstein-Price, Easom, and Schwefel*. The objective was to find a zone of data with maximum similarity between each test function and a quadratic function, which is presented below. The optimization problem was executed using two different initial solutions for each function, from which the best results were selected.
*More information about these functions can be found in Pohlheim (2006) and (Surjanovic and Bingham, 2013)

    i.   Typical optimization test functions in literature were chosen as the **functions to approximate** as previously mentioned.

    ii.   The **function to superimpose** was a metamodel; a second order polynomial regression of the form: $Z = f(x_1, x_2) = \beta_0 + \beta_1 x_1^2 + \beta_2 x_2^2$. The ranges of the variables of this quadratic function were the same in which each test function varies respectively.

    iii.   To generate the **grid of experimental points**, the ranges of the variables of each test function were divided according to a specific value of delta x (Δx) as indicated in Table 2.

    iv.   The **optimization problem**, as given by (1), also had to find metamodel parameter estimates β_0, β_1 and β_2. An additional constraint to this LP formulation included a minimum and maximum value for each of the estimates to vary in of -1000 and 1000. The parameters for the Solver included:

        • The use of multiple starting points using a population size of 100.

        • A convergence of $1\text{x}10^{-4}$.

        • Bounds are required on variables.

        • A level of precision of $1\text{x}10^{-3}$ for the functions Sphere, Rosenbrock, Griewank, Goldstein-Price, Easom, and Schwefel; a level of precision of $1\text{x}10^{-9}$ for the function Rastrigin.
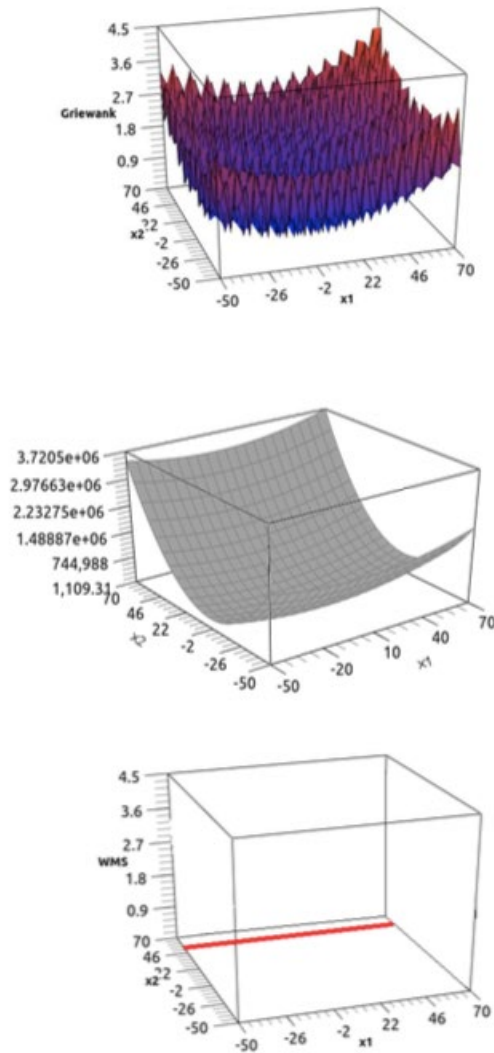
Figure 4: Griewank function (s variables): $f(X) = \sum_{r=1}^{s} \frac{x_r^2}{4000} - \prod_{r=1}^{s} \cos\left(\frac{x_r}{\sqrt{r}}\right) + 1, \ -50 \leq x_r \leq 70$

The best objective value found, SSE value and WMS are reported in Table 2 for each of the test functions.
Although estimated metamodel parameters are omitted to emphasize analysis on window results, the Sphere's, Griewank's, and Schwefel's quadratic function, $Z = f(\boldsymbol{R})$, followed the test function shape, as is shown in figure 4 for Griewank case. The WMS generated for functions Griewank and Schwefel potentially detected a zone of maximum similarity. The global solution for each global test function is additionally included. The solution is in all cases contained within the window of maximum similarity in, at least, one of the independent variables. For three other cases: Sphere, Griewank and Easom, the global solution is contained within the window for all the independent variables. The simplest case, the Sphere, is the most evident case where the quadratic function is a good descriptor of the 'data at hand', unlike the results of the remaining test functions which displayed quadratic functions of varied shapes and consequently their WMS were adjusted in varied zones.

Table 2: Results for typical optimization test functions in literature and metamodel

Function	x_1^L	x_1^U	x_2^L	x_2^U	SSE	Objective Function Value	WMS Dimensions	Global Solution	Is solution within window? (Y/N)
Sphere ($\Delta x =1.024$)	-5.12	5.12	-5.12	5.12	1.70E-10	**-0.70**	10.24 x 10.24	(0,0)	**Y**
Rosenbrock ($\Delta x = 0.5$)	-2	5	-2	-1.5	0	-0.36	7 x 0.50	(1,1)	N
Rastrigin ($\Delta x = 0.5$)	4.50	5	-5	5	0	-0.41	0.50 x 10	(0,0)	N
Griewank ($\Delta x = 2$)	-50	70	32	34	0	**-0.85**	120 x 2.00	(0,0)	**Y**
Goldstein-Price ($\Delta x = 0.25$)	-2	2	-0.25	0	0	-0.27	4 x 0.25	(0, -1)	N
Easom ($\Delta x= pi$)	-35.15	31.55	-35.79	66.64	5.71	**-1.01**	66.71 x 102.43	(0,0)	**Y**
Schwefel ($\Delta x = 5$)	495	499.52	-500	500	0	-1.26	4.92 x 1000	(420.97,420.97)	N

5. Conclusion and Future Work

This work proposed the use of WMS for future optimization by similarity. The method intends to find the experimental region where a model with desirable characteristics is a good descriptor of the data at hand. An evaluation case using function AOG 1 was presented. According to these results, the method demonstrates the potential to find regions of similarity between two responses where optimality can be a pattern of interest.

The evaluations of the method in seven unconstrained global optimization test functions served to show the use of window of maximum similarity in examples of functions with different shapes. Also, it was observed in the evaluations that the WMS method potentially detected zones of maximum similarity between the different test functions and a quadratic function.

According to these results, the method demonstrates the potential to find regions of similarity between two responses where optimality can be a pattern of interest and can be a useful tool for exploration of simulated data to find, at least a local optimum.

In many cases, the WMS obtained by the method were limited to take the minimum size or epsilon value assigned, which is why future work includes:

1. Substitute Optimization of one composite objective function to multiple criterion optimization, as presented in Rodríguez-Yáñez, Méndez-Vázquez, and Cabrera-Ríos, M., 2014.
2. Include more variables to the test functions to evaluate.
3. Experiment using alternative metamodels.
4. Use Design of Experiments for Experimental Region

References

Barton, R.R., Issues in development of simultaneous forward-inverse metamodels, *Proceedings of the 2005 Winter Simulation Conference*, pp. 209-217, 2005.

Couckuyt, I., Gorissen, D., DeTurck, F and Dhaene, T., Inverse surrogate modeling: output performance space sampling, *13th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, 10.2514/6.2010-9393, 2010.

Frontline Systems, Inc., Standard excel solver - GRG nonlinear solver stopping condition, Available: http://www.solver.com/standard-excel-solver-grg-nonlinear-solver-stopping-conditions, March 5, 2015.

Montgomery, D.C., *Design and Analysis of Experiments*, 2nd Edition, Limusa Wiley, 2010.

Pohlheim, H., GEATbx: Genetic and evolutionary algorithm toolbox for use with Matlab documentation, version 3.8. Available: http://www.geatbx.com/docu/fcnindex-01.html, February 12, 2014.

Rivera-Nazario, J.C., and Cabrera-Ríos, M., Searching the experimental region where a metamodel fits best, *Proceedings of the 2013 Industrial and Systems Engineering Research Conference,* 2013.

Rodríguez-Yáñez, A. B., Méndez-Vázquez, Y., and Cabrera-Ríos, M., Simulation-based process windows simultaneously considering two and three conflicting criteria in injection molding, *Production & manufacturing research*, 2(1), 603–623. doi:10.1080/21693277.2014.949359, 2014.

Surjanovic, S., and Bingham, D., Virtual library of simulation experiments: test functions and datasets, Available: http://www.sfu.ca/~ssurjano, May 21, 2015.

Bibliography

Veronica Diaz Pacheco is a graduate student, researcher and teaching assistant to the Work Measurement course at University of Puerto Rico, Mayagüez. She is a member of the Applied Optimization Group at UPRM. Her research interests include optimization, simulation, operations research, and modeling.

Mauricio Cabrera-Ríos is a Professor in the Industrial Engineering Department at University of Puerto Rico at Mayagüez. He obtained his MS and PhD degrees from The Ohio State University in 1999 and 2002 respectively. His interests include characterization, modeling and optimization of manufacturing and biological systems. He leads the Applied Optimization Group and is the editor-in-chief of JOUST, the Journal of Undergraduate Research Students. Mauricio.cabrera1@upr.edu, Facebook: AppliedOptUPRM

Mary Carmen Acosta Cervantes obtained the MS in Systems Engineering from the Autonomous University of Nuevo León, México in 2012 and the MS in Industrial Engineering from University of Puerto Rico at Mayagüez in 2015. Her interests include mathematical modeling and operations research. Marycarmen8070@gmail.com