Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

RESEARCH ARTICLE

Towards optimization by matching of response surfaces: Finding windows of maximal similarity

Verónica Díaz | Mary Acosta | Mauricio Cabrera-Rios

The Applied Optimization Group, Industrial Engineering Department, University of Puerto Rico – Mayagüez, Mayagüez, Puerto Rico, USA.

Correspondence
Mauricio Cabrera-Rios, Department of Industrial Engineering Department, University of Puerto Rico - Mayagüez, Mayagüez, 00682, Puerto Rico.
Email: Mauricio.cabrera1@upr.edu

Funding Information
USDA-NIFA Be Aware Project
Grant/Award Number: 2015-38422-24064 sub award 1000000920

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 functions' input space where this resemblance occurs, we call a Window of Maximal 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, that include several typical global optimization test functions in literature, are presented in order to demonstrate the method's feasibility and capability for generating a two-dimensional WMS. This tool is a viable element that will serve for the future development of Optimization by Similarity.

KEYWORDS

Metamodels, Optimization, Response Surface Methodology, Simulation

1 | INTRODUCTION

Response surfaces are typically used for optimization since they provide a visual estimate of the behavior of a function across its input space. Their employment for this purpose implies a forward mapping in the sense that, after data is sampled and an equation that can best approximate it is constructed, an experimental procedure is realized until an attractive solution -hopefully an optimum- is found. As an alternative, in this work the idea of inverse mappings is explored towards the optimization end, where desired output characteristics are associated to a specific region in a function's input space.

Consider the following hypothetical example of fitting a linear regression to predict student weight based on their height for a sample of size 7 (Please see Figure 1). As can be noted from the graph, there are some points (regions) where the line is a better predictor than others. The points in black show where data 'behaves' in a desirable manner. Note that desirable behavior is defined as that region where the data is most similar to a particular function, in this case, to the line that minimizes the sum of squared errors. Since this phenomenon can occur whenever

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

using modeling to approximate datasets, we propose extending the concept of desirable behavior - output characteristics - to include regions in datasets where it could ‘look like’ a function that has optimality properties; if we could identify a region in the dataset where the it is most similar to a function with optimality properties, we would have had found an area of potential optimality. But how may two models generated from the same data be compared?

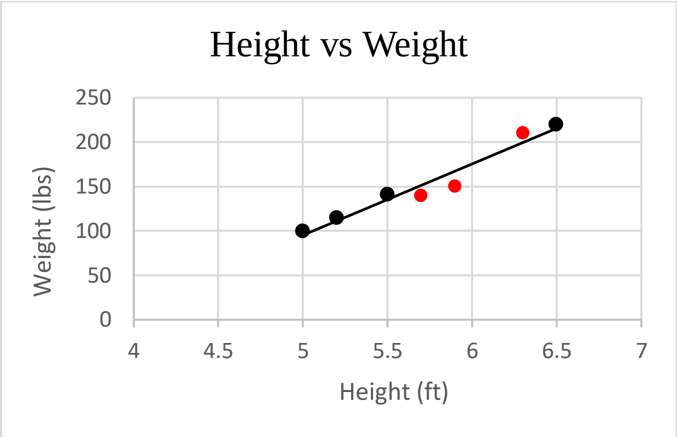

Figure 1 Hypothetical regression example, points in black show regions where model is better predictor.

Metamodeling is when a complex model, like the ones frequently used for simulations, is approximated by another, typically simplified, one. Metamodels are often used for optimization purposes and require forward mapping procedures to experimentally find optimality conditions for a particular process. It is also common that a metamodel’s parameters are estimated via minimization of an error function until the most competitive fit is found. In this work, metamodels are fitted in a different manner but also towards an optimization end; since we are searching for the region of maximum resemblance between a data generating model and a metamodel with optimality properties, the parameter estimates will differ. By finding a region of maximal similarity, we are looking to generate an inverse map in order identify a window in the input space where optimality may be present.

Inverse mappings, when a function’s input is a specific desired performance and its output its associated controllable variable settings, was approached in [3]. From the intricacies the author mentions, it was noted that the task of inverse mappings is often reduced to finding one (or more) input parameter combinations for only 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 [4]. 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 [6], our Optimization by Similarity method aims 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 (Please

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

refer to Figure 2). 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*.

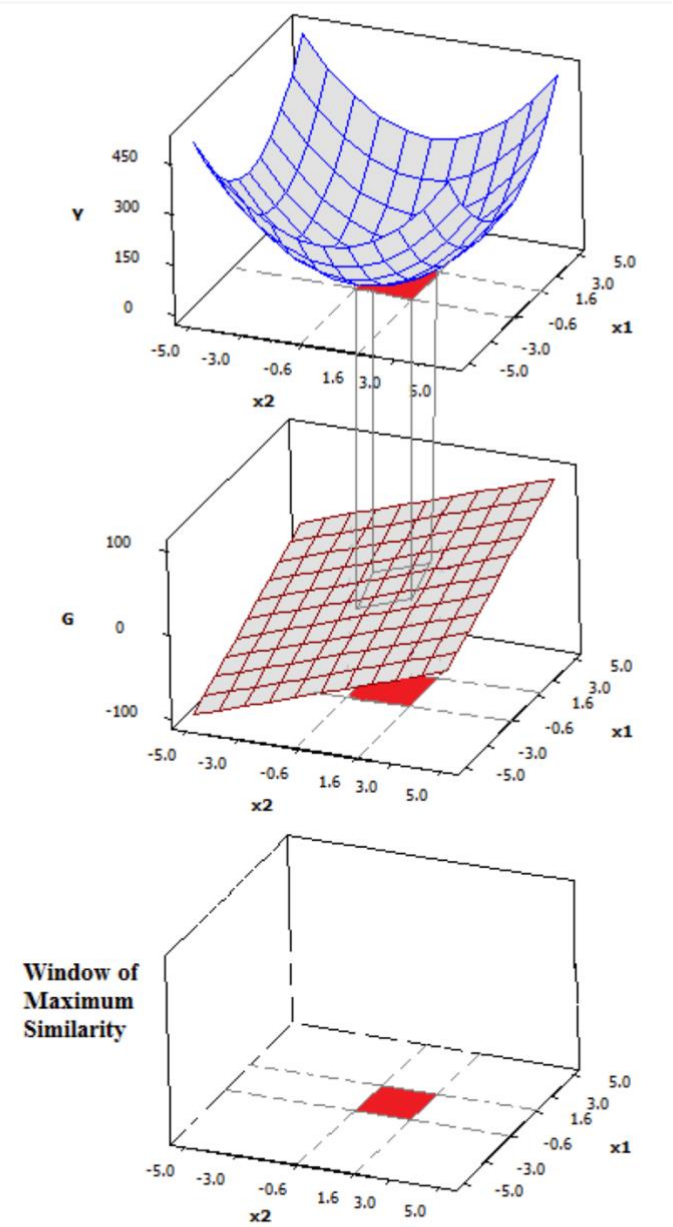

Figure 2 Example for construction of the Optimization by Matching of Responses where function to approximate data is quadratic and function to superimpose is linear; region in red represents Window of Maximal Similarity.

2 | TECNIQUES USED IN THIS WORK

2.1 | LEAST SQUARES METHOD

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;

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

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

In this work, the SSE is given by:

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

Where , $\boldsymbol{Y} = \boldsymbol{f}(\boldsymbol{R})$, is the response of the function to approximate that needs to be optimally addressed and $\boldsymbol{Z} = \boldsymbol{f}(\boldsymbol{R})$, 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 size, Δx can be chosen when the input variable initialization values are selected not to be integers. This step size 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 [7].

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, $\boldsymbol{Z} = \boldsymbol{f}(\boldsymbol{R})$, could be fixed 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 (2) which minimizes the functions' SSE. The solution will be the location and size of a window in the experimental region where the maximum similarity between the responses occurs, except when Z is a metamodel in which case the solution includes the metamodel parameter estimates. The optimization problem formulation and an example are illustrated below:

Find

$$x_i^L, x_i^U, to$$

Min

$$\frac{1}{3}\left([log\,(SSE+1)] - [log(|x_1^U - x_1^L| + 1)] - [log(|x_2^U - x_2^L| + 1)] \right) \qquad (2)$$

Subject to

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

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

* Constraints for experimental region bounds, respectively.
**For the integer variable cases, add to the formulation the following restriction:

x_i^L, x_i^U = integer

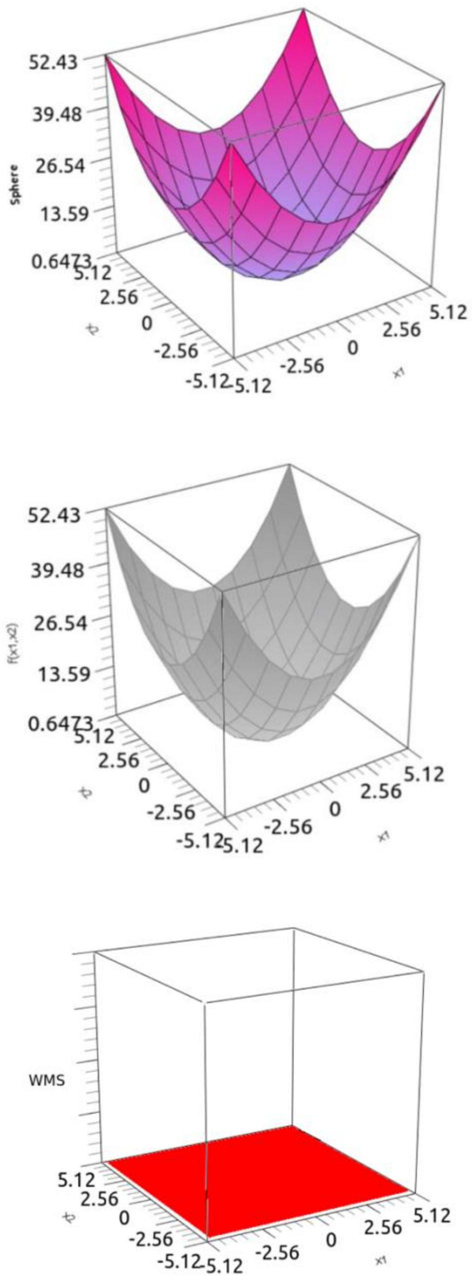

Figure 3 Example of Window of Maximal Similarity found between Sphere and second order polynomial regression.

In Figure 3, an example is presented to demonstrate the application of the proposed method, where function to approximate, $\boldsymbol{Y} = \boldsymbol{f}(\boldsymbol{R})$, is the sphere and the function to superimpose, $\boldsymbol{Z} = \boldsymbol{f}(\boldsymbol{R})$, is a second order polynomial regression of the form:

$$Z = f(x_1, x_2) = \beta_0 + \beta_1 x_1^2 + \beta_2 x_2^2 \quad (3)$$

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

Since the function to superimpose is a metamodel, the optimization model had to additionally ***find*** metamodel parameters $\boldsymbol{\beta_0}, \boldsymbol{\beta_1}, \ldots \boldsymbol{\beta_n}$. For the special case where the two functions have the exact same shape, as was the case presented for Figure 3, the problem will evidently have infinitely many solutions since the window of maximum similarity could adjust itself in any range in the experimental region.

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 ($\boldsymbol{x_i^U - x_i^L}$) are present:

1. in the objective function in order to avoid window size dependency on parameters
2. in the constraints because a minimum value had to be included for the formulation to be successful and not contain a single point. This suggested value, $\boldsymbol{\varepsilon = 1 \times 10^{-6}}$, is known as the non-archimedean constant, a value commonly used for computational purposes.

Additional constraints include a range where to define the window's upper and lower bounds, in accordance to the span in which each variable varies.

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 [7].

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**.**

1. 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 $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ were [-5, 5]. *AOG_1* is given by:

$$Y = f(x_1, x_2) = \begin{cases} 5x_1^2 + 5x_2^2 \quad if \;\; x_1 \in [-2, 1], x_2 \in [-3, 0] \\ 500 - 5x_1 + 5x_2 \;\; for\ any\ other\ point \end{cases} \tag{4}$$

2. The **function to superimpose** was fixed for this case and given by

$$\boldsymbol{Z = f(x_1, x_2) = 5x_1^2 + 5x_2^2} \tag{5}$$

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

The ranges $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$were also [-5, 5].

3. The **experimental region** is a grid that contains 121 points; $\boldsymbol{x_1} \in [\mathbf{5}, \mathbf{5}], \boldsymbol{x_2} \in [\mathbf{5}, \mathbf{5}]$.
4. The **optimization problem formulation** is given by (2).
5. The last step is to **optimize the model**. The global minimum for *AOG_1* is given by solution (0, 0), with a corresponding objective value of 0. Essentially, the maximum similarity would be found if the WMS was 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 $\mathbf{1} \times \mathbf{10^{-7}}$.
    - A convergence of $\mathbf{1} \times \mathbf{10^{-4}}$.

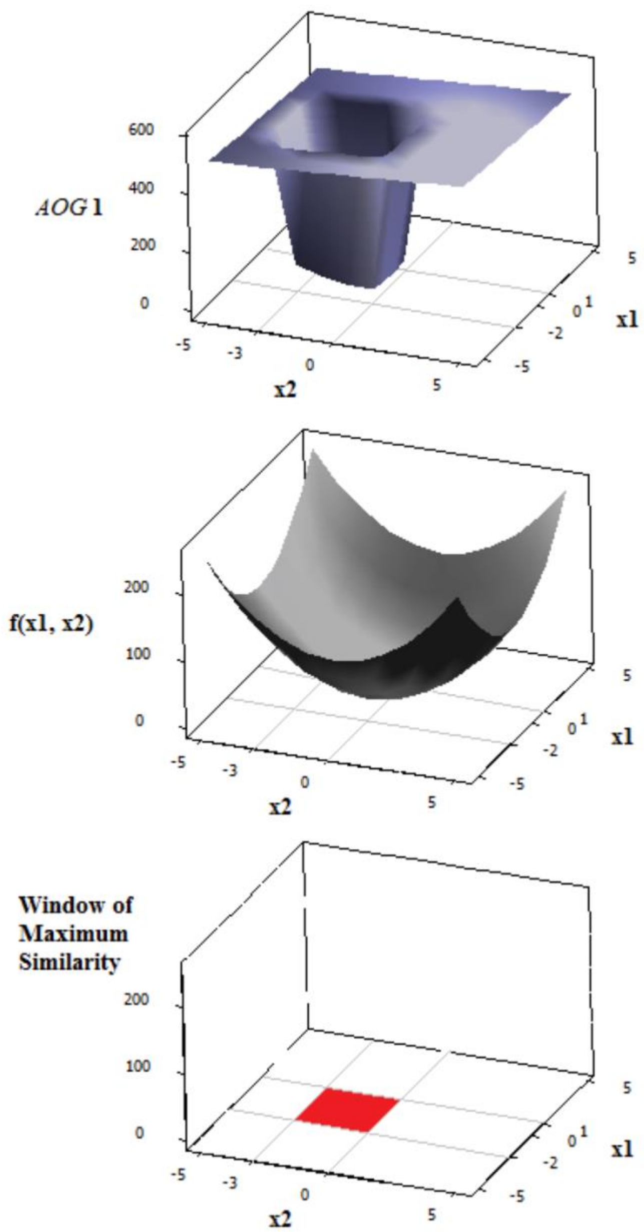

Figure 4 Window of Maximum Similarity between *AOG_1* and (static) quadratic function.

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

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 4. 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 Results for *AOG_1*, and quadratic (static) function to superimpose.

ToV	Best Solution (WMS Location)				Best Objective Value Found				WMS Size
	x_1^L	x_1^U	x_2^L	x_2^U	*OFV*	*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

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.

*For more information, please refer to [2,8].

1. **Function(s) to approximate** were typical optimization test functions in literature, as previously mentioned.
2. The **function to superimpose** was a metamodel, a second order polynomial linear regression of the form

$$Z = f(x_1, x_2) = \beta_0 + \beta_1 x_1^2 + \beta_2 x_2^2 \,. \tag{6}$$

   The ranges of the variable bounds for this quadratic function were the same in which test function varies respectively.
3. 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) (see Table 2).

Table 2 Step size used to generate grid of experimental points for global optimization test functions.

Function	Δx
Sphere	1.024
Rosenbrock	0.5
Rastrigin	0.5

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

Griewank	2
Goldstein Price	0.25
Easom	pi
Schwefel	5

4. The **optimization problem formulation**, as given by (2), also had to ***find*** metamodel parameters, β_0, β_1 and β_2. An additional constraint to this linear programming formulation included a minimum and maximum value for each of the parameter estimates to vary in of [-1000, 1000]. The parameters for the Solver included:
    - The use of multiple starting points using a population of size 100.
    - A convergence of $1\text{x}10^{-4}$.
    - Bounds required on variables.
    - A level of precision of 1×10^{-3} for the functions Sphere, Rosenbrock, Griewank, Goldstein-Price, Easom, and Schwefel; a level of precision of 1×10^{-9} for the function Rastrigin.

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

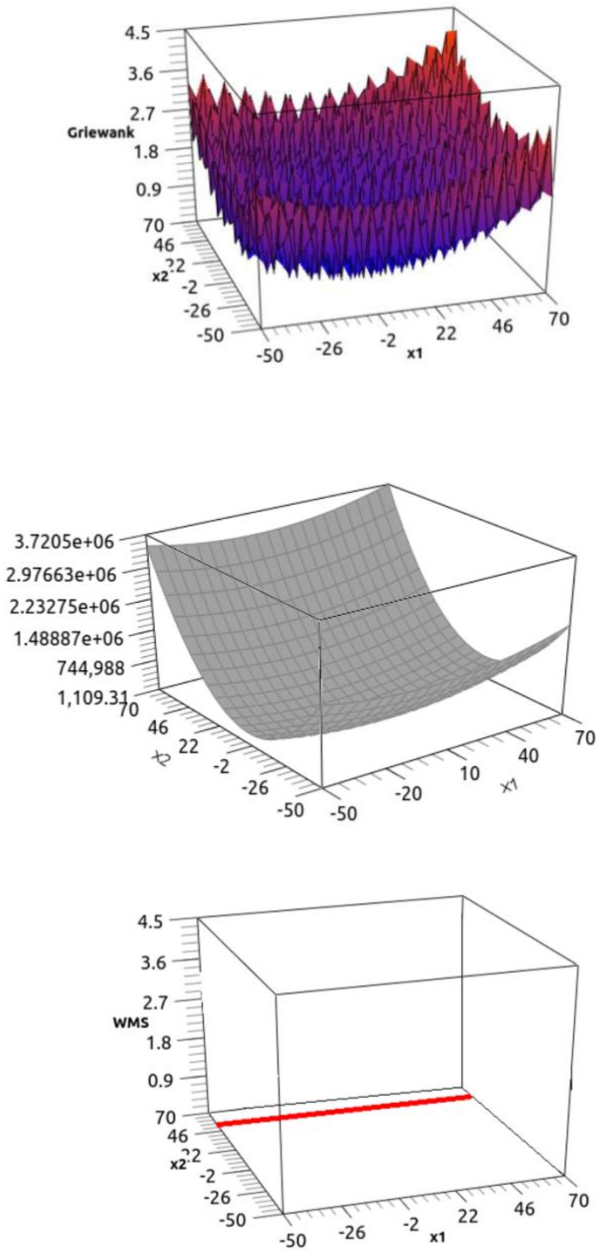

Figure 5 Griewank (s variables), given by: $f(X) = \sum_{r=1}^{s} \frac{x_r^2}{4000} - \prod_{r=1}^{s} \cos\left(\frac{x_r}{\sqrt{r}}\right) + 1,\ -50 \leq x_r \leq 70$, and quadratic function.

The best objective value found, SSE value and WMS are reported in Table 3 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, $\boldsymbol{Z} = \boldsymbol{f}(\boldsymbol{R})$, followed the test function shape, as is shown in Figure 5 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

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

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 3 Results for optimization test functions in literature and (movable) metamodel.

Function	Best solution (WMS location)				Global solution	Solution within window?	*Best OFV*	WMS Size
	x_1^L	x_1^U	x_2^L	x_2^U				
Sphere	-5.12	5.12	-5.12	5.12	(0,0)	**Y**	**-0.70**	10.24 x 10.24
Rosenbrock	-2	5	-2	-1.5	(1,1)	N	-0.36	7 x 0.50
Rastrigin	4.50	5	-5	5	(0,0)	N	-0.41	0.50 x 10
Griewank	-50	70	32	34	(0,0)	**Y**	**-0.85**	120 x 2.00
Goldstein-Price	-2	2	-0.25	0	(0, -1)	N	-0.27	4 x 0.25
Easom	-35.15	31.55	-35.79	66.64	(0,0)	**Y**	**-1.01**	66.71 x 102.43
Schwefel	495	499.52	-500	500	(420.97, 420.97)	N	-1.26	4.92 x 1000

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.

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

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 single composite objective by multiple criterion optimization, as presented in [5].
2. Include more variables to evaluate test functions on.
3. Experiment using alternative metamodels with different forms.
4. Use design of experiment to sample from the experimental region.

AKNOWLEDGEMENTS

This work was supported by **USDA-NIFA** Award 2015-38422-24064 sub award 1000000920. BE AWARE Project.

CONFLICT OF INTEREST

Authors have no conflict of interest relevant to this article.

REFERENCES

[1] D. C. Motgomery, Design and analysis of Experiments, Second Edition, Limusa Wiley, 2010.

[2] H. Pohlheim, “GEATbx: Genetic and evolutionary algorithm toolbox for use tith MATLAB documentation,” [Online]. Available: http://www.geatbx.com/docu/fcnindex-01.html. [Accessed May 2019].

[3] R. Barton, “Issues in development of simultaneous forward-inverse metamodels,” in *37th Winter Simulation Conference*, Orlando, 2005.

[4] I. Couckuyt, D. Gorissen, T. Dhaene and F. DeTurck, “Inverse surrogate modeling: output performance space sampling,” in *13th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Fort Worth, 2010. https://doi.org/10.2514/6.2010-9393.

[5] A. B. Rodriguez-Yañez, Y. Méndez-Vázquez and M. Cabrera-Rios, “Simulation-based process windows simultaneously considering two and three conflicting criteria in injection molding,” *Production & manufacturing research,* pp. 603-623. https://doi.org/10.1080/21693277.2014.949359, 2014.

[6] Frontline Systems Inc., “Standard Excel solver - GRG nonlinear solver stopping condition,” [Online]. Available: http://www.solver.com/standard-excel-solver-grg-nonlinear-solver-stopping-conditions.. [Accessed May 2019].

[7] S. Surjanovic and D. Bingham, “Virtual library of simulation experiments: test functions anddatasets,” [Online]. Available: http://www.sfu.ca/~ssurjano. [Accessed May 2019].

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

[8] J. C. Rivera-Nazario and M. Cabrera-Rios, “Searching the experimental region where a metamodel fits best,” in *IIE Annual Conference Proceedings*, 2013. Available: https://uprm.idm.oclc.org/docview/1471960482?accountid=28498.