SANDIA REPORT
SAND2021-15040
Printed November 2021

Sandia National Laboratories

Estimating the Adequacy of a Multi-Objective Optimization

Lucas Waddell, John H. Gauthier, Matthew J. Hoffman, Denise D. Padilla, Stephen M. Henry, Alexander I. Dessanti, Adam J. Pierson

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico
87185 and Livermore,
California 94550

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology & Engineering Solutions of Sandia, LLC.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@osti.gov
Online ordering: http://www.osti.gov/scitech

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.gov
Online order: https://classic.ntis.gov/help/order-methods/

ABSTRACT

Multi-objective optimization methods can be criticized for lacking a statistically valid measure of the quality and representativeness of a solution. This stance is especially relevant to metaheuristic optimization approaches but can also apply to other methods that typically might only report a small representative subset of a Pareto frontier. Here we present a method to address this deficiency based on random sampling of a solution space to determine, with a specified level of confidence, the fraction of the solution space that is surpassed by an optimization. The Superiority of Multi-Objective Optimization to Random Sampling, or SMORS method, can evaluate quality and representativeness using dominance or other measures, e.g., a spacing measure for high-dimensional spaces. SMORS has been tested in a combinatorial optimization context using a genetic algorithm but could be useful for other optimization methods.

ACKNOWLEDGEMENTS

The authors would like to thank Stanislaw Ulam for inspiring this work. We would also like to thank Garrett Dranichak for his thoughtful review. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

This page left blank

DEFINITIONS

Term	Definition
Confidence	Probability that an estimated value is unlikely to be wrong (here with reference to the estimate of superiority to randomness being exaggerated or overestimated)
Dimension	An axis in a multi-objective optimization defined by one objective or combination of objectives, here with reference to a low-dimension or high-dimension problem
Dominated	A solution in an optimization solution set or during the progression of an optimization computation for which some other solution in the solution set or intermediate set is equal to or better than the first solution in all objectives and strictly better than it in at least one objective; often applied to determine solutions that should be omitted from further consideration during the progression of an evolutionary algorithm
Expanse	Distance between the minimum and maximum points in an optimization dimension
High-dimension problem	Multi-objective optimization problem typically encompassing greater than eight dimensions
Low-dimension problem	Multi-objective optimization problem typically encompassing less than about four dimensions
MILP	Mixed-integer linear program; a mathematical optimization problem
Non-dominated	A solution in an optimization solution set or during the progression of an optimization computation for which no other solution in the solution set or intermediate set is better in all values calculated for every optimization dimension; often applied transiently to determine what solutions should be used during the progression of an evolutionary algorithm
Pareto frontier	The exact solution set to a multi-objective optimization; the set of non-dominated solutions
Point	An individual solution within the optimization solution space
Representative Pareto frontier	Approximation of a Pareto frontier, typically not necessarily optimal and involving a subset of actual Pareto frontier (also sometimes called a representative solution set)
Search space	Domain of the function to be optimized; the feasible region of the set of all possible solutions
SMORS	Superiority of Multi-Objective Optimization to Random Sampling
Spacing	Distance between points in a representative multi-objective optimization solution set; often applied transiently to determine what points should be used during the progression of an evolutionary algorithm (e.g., niching)
Solution set	The result of optimization; for a single-objective optimization, a single solution; for a multi-objective optimization, a collection of solutions
Solution space	Search space
Superiority to randomness	The distance, in units of probability, of a computed, representative Pareto frontier from a random sampling of the solution space

Term	Definition
Very high-dimension problem	Multi-objective optimization problem encompassing significantly more dimensions than a high-dimension problem, e.g., 20 dimensions

1. MOTIVATION

A trend in stakeholder-driven optimizations is toward higher-dimensional, multi-objective analyses that often require demanding evaluations. (See Definitions for High-dimension problem. Demanding evaluations are here defined as computation times of all required evaluations exceeding a day.) The reason is that these stakeholders are interested in understanding the space of the best decisions more than simply knowing the single optimal, and perhaps brittle, solution. In this environment, adequacy of an analysis must be assured: was a statistically valid, representative set of the best solutions identified? The assurance is especially necessary for metaheuristic optimization methods (evolutionary algorithms) but could also apply to mathematical optimization approaches (e.g., mixed-integer linear programming). Consider, for example:

1. multi-objective optimizations based on a grid of optimization points: are we missing significant behavior between the grid nodes, perhaps behavior that dominates the solutions at the grid nodes?
2. high-dimensional, multi-objective optimization with expensive evaluations: how much effort should be spent in computing each solution and when are the results adequate?
3. very high-dimensional, multi-objective optimizations where the goal of the optimization is not to find optimal solutions (virtually any randomly chosen solution is non-dominated), but rather to find representative coverage of the solution space: is the solution set well-spaced, does it contain the diversity of possibilities, and does it reflect the density of solutions in the Pareto frontier?
4. metaheuristic multi-objective optimization that uses a discretization of the decision-variable values: is the discretization used adequate?

To address these issues, we require a general, statistically valid approach to measure our confidence in the quality of an optimization and thus the quality of the decisions supported by the optimization.

2. BACKGROUND

The absence of optimality conditions and the absence of quantitative confidence measures in metaheuristic-derived solutions have been addressed by various methods. One technique is to quantify the quality with which an algorithm explores the feasible search space. See, for example,

- the one-dimensional search space visualization tool (Mach and Zetakova 2002),
- the two-dimensional "coverage maps" used to visualize search space coverage and convergence behavior (Shine and Eick 1997), and
- the clustering-based "coverage of the relevant search space" metric (Wehrens et al. 1998).

Another approach in the literature has been to assert the quality of a metaheuristic for a specific application by comparing the final solution to that obtained by a random search algorithm (see, for instance, Marrison and Stengel 1994, Godefroid and Khurshid 2004). For a related problem, Alyahya and Rowe (2016), use random sampling to estimate the number of local optima, to aid in commenting on the quality of found local optima or the assurance that a global optimum has been seen.

The SMORS method, presented in this report, combines these two previously disconnected approaches by comparing optimization-generated solution sets to randomly generated solutions. The goal of the SMORS method is not simply to claim the optimization's superiority, but to additionally provide a quantitative statistical bound on the quality of the exploration and coverage of the feasible search space.

3. THE SMORS METHOD

The motivation for the SMORS method comes from an analogy to the Monte Carlo method, as applied to integration (see, for instance, Caflisch 1998). As shown in Figure 1a, the Monte Carlo method can be used to generate random points on a bounded plane, and the ratio of the number of points falling below a curve to the total number of points generated, times the area of the bounded region, is an estimate of the area under the curve. Similarly, a Pareto frontier calculated by a multi-objective optimization can be considered to act like a curve within a solution space. In Figure 1b, points (individual solutions to an optimization problem) are randomly generated within the feasible solution space. Every random point is deemed to be either inferior or superior to the multi-objective solution set. In Section 3.2, we will discuss this designation in greater detail, but for now, we can think of a superior random point as one that dominates at least one solution in the multi-objective solution set. An inferior random point is simply one that is not superior to the multi-objective solution set.

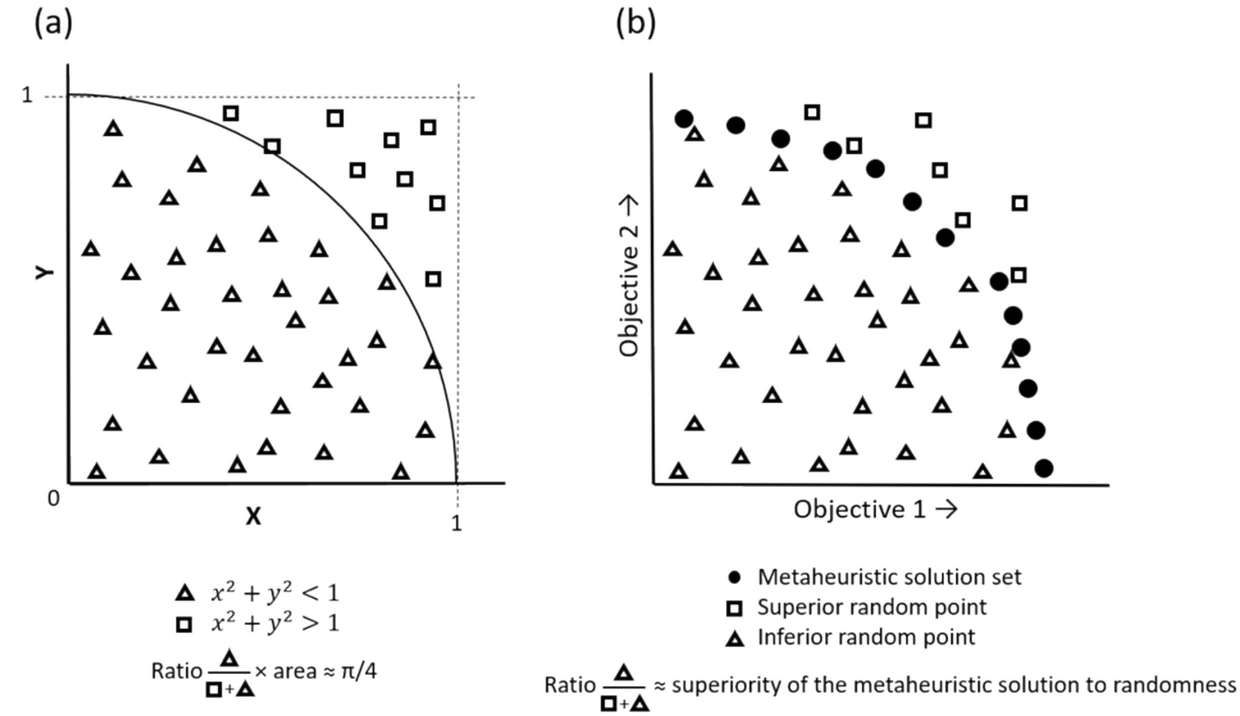

Figure 1 (a) Estimate of π/4 using the Monte Carlo method. (b) Analogous method for estimating the "superiority to randomness" of a multi-objective solution set.

The ratio between the number of randomly generated points that are inferior to the multi-objective solution set and the total number of randomly generated points is calculated. This ratio is an estimate of the probability that a randomly selected solution will be inferior to the multi-objective solution set. For example, we could use the random sample pictured in Figure 1b to estimate this probability to be about 84%. We will refer to this ratio as the multi-objective solution set's estimated "superiority to randomness."

We would, of course, like this "superiority to randomness" ratio to be as large as possible. In fact, the SMORS method is based on the special case where the random sample results in a ratio that is equal to 1. In other words, the SMORS method involves taking random samples, and drawing statistical conclusions when none of these random points are superior to the multi-objective solution set.

The key insight is that each randomly sampled point can be thought of as a trial from a binomial experiment with an unknown probability of success θ, where a trial is successful if the random point is inferior to the multi-objective solution set. In this sense, θ represents the true proportion of the entire search space that is inferior to the multi-objective solution set (i.e., the multi-objective solution set's true superiority to randomness). We can use the Clopper-Pearson confidence interval for binomial distributions (Clopper and Pearson 1934) to compute a confidence interval for θ.

For the special case where every observed trial is a success (i.e., all sampled points are inferior to the multi-objective solution set), the 1-α one-sided Clopper-Pearson confidence interval for the population proportion of successes θ is given by

$$(\alpha^{1/n},1),$$

as described in (Thulin 2014), where *n* is the number of trials.

Suppose we would like to claim that a given proportion, say $\hat{\theta}$, of the search space is inferior to the multi-objective solution set. We can use the confidence interval to calculate the number of random samples without seeing a superior point that are necessary to assert (with confidence $1-\alpha$) that the true proportion of the search space that is inferior to our multi-objective solution set is *at least* $\hat{\theta}$. Specifically, the number of trials necessary is given by

$$n = \lceil \ln(\alpha)/\ln(\hat{\theta}) \rceil,$$

which results in a $1-\alpha$ confidence interval of $(\hat{\theta},1)$ for the true proportion of the search space that is inferior to the multi-objective solution set.

For example, we would need to sample *n* = 298 random points without seeing a point superior to the multi-objective solution set to be 95% confident that the multi-objective solution set's true superiority to randomness is at least 0.99. In other words, we would be 95% confident that at least 99% of the points in the search space are inferior to the multi-objective solution set. Put another way, we would be 95% confident that *at most* 1% of the of the points in the search space are superior to the multi-objective solution set.

Table 1 offers numbers of random samples needed—without finding a superior point—to achieve various confidence levels of various lower bounds on the multi-objective solution set's true superiority to randomness.

Table 1 Number of random trials (n), without seeing a superior point, required to achieve various confidence levels ($1-\alpha$) for several lower bounds ($\hat{\theta}$) on the superiority to randomness of a multi-objective solution set.

Superiority-to-Randomness Fraction	Confidence = 95%	Confidence = 99%	Confidence = 99.9%
0.99	298	458	687
0.999	2,994	4,603	6,904
0.9999	29,956	46,049	69,074

3.1. Using the SMORS method on a real-world problem

The U.S. Army is considering development of the Squad Multi-purpose Equipment Transport (SMET), a semi-autonomous mobile platform for carrying gear for an infantry squad during operations, i.e., a robotic pack mule. SMET has a number of performance objectives in addition to carrying cargo: rapid battery recharge, mission payload hosting, reliable, survivable, etc. SMET must also be cost effective to buy, to operate, and to sustain.

Before the Army requested proposals for SMET, a trade-space study was conducted to determine what such vehicles might look like. The basic component areas of SMET included powertrain, chassis and body, communications and control, and sensors. Each of these areas had dozens of possible technology options, which when combined gave an approximate solution space size of 2.638×10^{22} (dependencies between some components, e.g., power supply and motor type, reduces this number somewhat). A multi-objective optimization using the Whole System Trades Analysis Tool (WSTAT) selected the best trade space to consider. (For more information about WSTAT and one of its applications, see Henry et al. 2016.) The genetic algorithm evaluated 1,063,323 configurations—fewer than 10^{-15} of all possible solutions. To build confidence in the solution, convergence was monitored, and statistics were kept on how often each technology option was tried. Although reasonable, these measures only indirectly indicated the quality of the SMET solution.

The SMORS method was applied to the SMET solution to determine, with 95% confidence, that the solution is at least in the upper 99.9% of the possible solutions. In other words, the method was used to provide 95% confidence that $0.999*2.638\times10^{22} = 2.635\times10^{22}$ solutions in the possible solution space were inferior to the WSTAT-computed SMET solution. (Two caveats: first, the 2.638×10^{19} solutions that comprise the remaining 0.1% of the solution space is still an enormous number of solutions; second, SMORS does not indicate how much better the SMET solution is to the 2.635×10^{22} inferior solutions, simply that it is better.) SMORS requires 2,994 random samples to determine the desired quality specification. In WSTAT, SMORS takes slightly over 7 minutes to build the random configurations and compare each for dominance against the SMET Pareto solutions. As no non-dominated random configurations were generated, it was possible to report to the Army customer that the SMET solution meets the desired 99.9% superiority with 95% confidence.

3.2. Determining the superiority of random points

Determination of random-solution superiority to the multi-objective solution set should typically use the same criteria that were used in the original generation of the multi-objective solution set. One way is to observe whether the point dominates (in the Pareto sense) any solutions in the multi-objective solution set. A solution dominates another if it performs better than or equal to that solution in all objectives and is strictly better in at least one objective. (As mentioned below, the dominance measure works best for problems with relatively few objectives.)

Other optimization criteria could be considered, such as whether the random point improves the spacing of the points in the multi-objective solution set or the expanse of the solution set (Dessanti et al. 2016). Figure 2 shows how a randomly generated point (indicated by "?") might improve the spacing of a multi-objective solution set and thus be considered superior, even though it does not dominate any of the solutions in the multi-objective solution set. In our context, the binomial distribution requires exactly two possible outcomes. Therefore, inferior random solutions are simply those that are not superior to the multi-objective solution set.

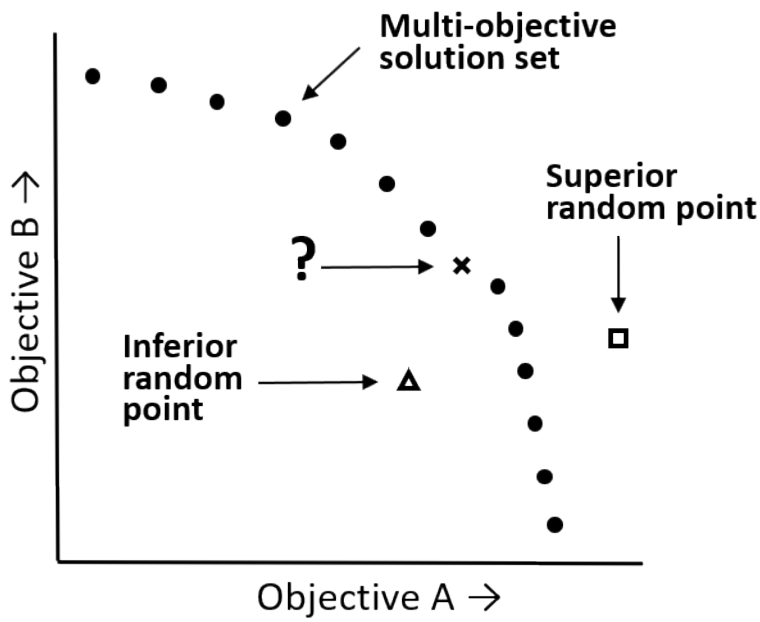

Figure 2 Illustration of a multi-objective solution set with a random inferior point, a random superior point, and a random point marked by "?" that, although it does not dominate any solutions in the multi-objective solution set, might or might not be considered superior depending on whether it improves the solution set's spacing.

In high-dimensional multi-objective optimization, where the number of dimensions is greater than approximately eight, the solution space is typically so large that most randomly generated points will be non-dominated (Ishibuchi 2008). In these cases, improved spacing and improved expanse could again be useful discriminators. Other techniques have been proposed for comparing points in high-dimensional spaces and could be considered for determining inferiority/superiority, e.g., *prefer* (Sulflow 2007), *favour* (Drechsler 2001), or reference-point-based non-dominated sorting (Deb and Jain 2014).

As a matter of course, we are only interested in generating feasible random points—those that satisfy all of the constraints of the optimization problem. Thus, in generating a random sample, infeasible points should be discarded or fixed in some unbiased manner—e.g., metaheuristic solution "healing" (Henry et al. 2015). Generating a representative sample from a highly constrained region can be challenging and is an active area of current research (see, for instance, Golchi and Loeppky 2016).

3.3. Using the SMORS method to develop a metaheuristic convergence criterion

Convergence criteria—numerical objectives of goodness that a solution must meet before algorithm termination—used in metaheuristic computations are often ad hoc, relying on an observed lack of progress in solution improvement (overviews of some of the most common metaheuristic stopping criteria can be found in Jain et al. 2001 and Saxena et al. 2016). The SMORS method could be used not only to judge the goodness of a metaheuristic solution set, but it could also be used before a metaheuristic solution set is computed to set a convergence criterion. For instance, suppose (as above) we want to be 95% confident that our metaheuristic solution set has at least 99% superiority to randomness. Then we could sample 298 random points independently of the metaheuristic computation and, as the metaheuristic optimization progresses, we could occasionally check to see if all of the random points are inferior to the current metaheuristic solution set. If so, the computation could cease, with the knowledge that the current solution set meets the initial numerical objective. This convergence criterion could be especially useful for solving high-dimensional, multi-objective optimizations with expensive evaluations.

3.4. Generality of the SMORS method

While the SMORS method provides the most value for optimization methods that lack formal optimality conditions or quantitative measures of solution quality (e.g., genetic algorithms, tabu search, simulated annealing), the method is general and could apply to any optimization technique. For instance, the SMORS method could also have value in determining, say, the coverage adequacy of a multi-objective solution calculated by a mixed integer linear program (MILP) method. In this case, although the adequacy of each individual solution can be assured by measuring the gap, the distribution of the solutions across the actual Pareto frontier could be insufficient and important regions of the Pareto frontier could be missed. (Indeed, one can imagine a case where a solution along a given MILP ray is dominated by a nearby solution along an untried ray.) Thus, the SMORS method can give quantitative bounds to the quality of MILP multi-objective solutions. For the SMORS method to be applicable, all that is needed is (1) a way to produce random individual solutions and (2) a way to determine whether a random solution is superior or inferior to a given optimization solution or solution set.

4. CONCLUSIONS

The SMORS method provides a quantitative lower bound of the probability that a randomly selected point from the search space will be inferior to a multi-objective solution set, along with a statistical measure of confidence in this bound. Although developed in the context of metaheuristics, the SMORS method can be applied to any optimization technique. We envision that a major application of the SMORS method will be to inspire confidence in multi-objective solutions when presenting results to stakeholders and decision makers.

If a random point is ever sampled that proves superior to a multi-objective solution set, this point can be added to the solution set. In the case of a metaheuristic optimization method, the computation can continue, thus increasing the quality of the metaheuristic solution set. In the case of a MILP optimization, the calculation can be adjusted for finer granularity.

Future work to be done in relation to the SMORS method includes 1) developing options for defining what it means for a random point to be inferior or superior for various classes of optimization problems, 2) researching techniques for generating random samples from highly constrained spaces, and 3) studying more effective techniques to visualize and present SMORS results to stakeholders.

REFERENCES

Alyahya, K., J.E. Rowe, Simple Random Sampling Estimation of the Number of Local Optima, in PPSN XIV 2016, LNCS 9921, J. Handl et al. (Eds.), 932–941, 2016. DOI: 10.1007/978-3-319-45823-6 87

Caflisch, R.E., Monte Carlo and quasi-Monte Carlo methods, *Acta Numerica*, Cambridge University Press, Vol. 7, 1−49, 1998.

Clopper, C.J., E.S. Pearson, The use of confidence or fiducial limits illustrated in the case of the binomial, Biometrika, 26, 404–413, 1934.

Deb, K., H. Jain, An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: Solving problems with box constraints, *IEEE Transactions on Evolutionary Computation*, 18(4), 577−601, 2014.

Dessanti, A.I., J.H. Gauthier, S.M. Henry, M.J. Hoffman, *Meta-Meta-Optimization for Integrated Requirements Development LDRD Final Report*, SAND2016-12620, Sandia Proprietary Information, Sandia National Laboratories, Albuquerque, NM, 2016.

Drechsler, N., R. Drechsler, B. Becker, Multi-objective optimisation based on relation favour, in *International Conference on Evolutionary Multi-Criterion Optimization*, 154−166, Springer Berlin Heidelberg, 2001.

Godefroid, P., S. Khurshid, Exploring very large state spaces using genetic algorithms, *International Journal on Software Tools for Technology Transfer*, 6(2), 117−127, 2004.

Golchi S., J. L. Loeppky, Monte Carlo based designs for constrained domains, arXiv:1512.07328v2, https://arxiv.org/abs/1512.07328, 2016.

Henry, S.M., G.S. Pankretz, L.A. Waddell, ``Healing the Achilles' Heel: Finding and Preserving Feasibility in Evolutionary Algorithms," SAND2015-9556 C, INFORMS Annual Meeting, Philadelphia, PA, Nov. 2015.

Henry, S.M., L.A. Waddell, M.R. DiNunzio, The Whole System Trades Analysis Tool for autonomous ground systems, *8th Annual Ground Vehicle Systems Engineering and Technology Symposium*, 2016.

Ishibuchi, H., N. Tsukamoto, Y. Nojima, Evolutionary many-objective optimization: A short review, in *IEEE Congress on Evolutionary Computation*, 2419−2426, 2008.

Jain, B.J., H. Pohlheim, J. Wegener, On termination criteria of evolutionary algorithms, *Proceedings of the Genetic and Evolutionary Computation Conference*, Morgan Kaufmann Publishers, 2001.

Mach, Z., M. Zetakova, Visualising genetic algorithms: A way through the labyrinth of search space, *Intelligent Technologies – Theory and Applications*, 279−285, 2002.

Marrison, C.I., R.F. Stengel, The use of random search and genetic algorithms to optimize stochastic robustness cost functions, *Proceedings of the American Control Conference*, 1994.

Saxena, D.K., A. Sinha, J.A. Duro, Q. Zhang, Entropy-based termination criterion for multiobjective genetic algorithms, *IEEE Transactions on Evolutionary Computation*, 20(4), 485−498, 2016.

Shine, W.B., C.F. Eick, Visualizing the evolution of genetic algorithm search processes, in *Proceedings of the IEEE International Conference on Evolutionary Computation*, 367–372, 1997.

DISTRIBUTION

Email—Internal

Name	Org.	Sandia Email Address
Mathew J. Hoffman	5523	mjhoffm@sandia.gov
Adam J. Pierson	5555	ajpiers@sandia.gov
Denise D. Padilla	5491	ddpadil@sandia.gov
Jason Brown	5493	jbrown2@sandia.gov
Alexander I. Dessanti	5493	adessan@sandia.gov
John H. Gauthier	5493	jhgauth@sandia.gov
Stephen M. Henry	5493	smhenry@sandia.gov
Technical Library	01977	sanddocs@sandia.gov

Email—External

Name	Company Email Address	Company Name
Lucas Waddell	lucas.waddell@bucknell.edu	Bucknell University

This page left blank

This page left blank

Sandia National Laboratories