—————-Response to reviewers———————

Multiple Surrogate Assisted Many-objective Optimization for
Computationally Expensive Engineering Design

Kalyan Shankar Bhattacharjee, Hemant Kumar Singh, Tapabrata Ray
School of Engineering and IT, The University of New South Wales, Canberra, Australia
Email: k.bhattacharjee@student.adfa.edu.au, {h.singh,t.ray}adfa.edu.au

We would like to thank all the reviewers for their constructive comments on our paper. We have undertaken a major revision of the manuscript based on all the reviewers' recommendations. Below, we respond to each of the comments and indicate changes that have been incorporated in the revised manuscript. Our responses in this document and the changes in the manuscript are highlighted in blue for an easy identification.

Reviewer 1

The paper discusses the solution of many-objective optimization (multi-objective optimization with more than three objectives) problems using surrogate models. The authors propose a solution method that utilize surrogate models. The proposed method is benchmarked with existing method from the literate and a vehicle design problem is used to illustrate the application and results of the presented approach.

The paper needs substantial changes regarding the presentation and the justification of the contributions. The contributions are not clear and there is insufficient discussion on what makes the proposed approach a contribution to the literature. Also, several important elements regarding surrogate modeling and their impacts on the proposed approach need more discussion in the paper to demonstrate the general applicability of the approach to a wider range of problems.

Here are more specific major comments for authors to consider:

- Surrogate modeling have been used in the optimization literature for a long time. Even in a single-objective optimization problem, multiple surrogate models are commonly used to represent both the objective and the constraints. I am having trouble with understanding what makes different for a multi-objective optimization problem. In either case, different surrogate models are necessary to represent different quantities and authors do not justify what makes a mere application of surrogate modeling to a multi-objective optimization problem. Authors do not propose a new way to do surrogate modeling. The only justification I see in the paper says: . their integration in an evolutionary framework is not straight-forward which essentially does not tell anything to the reader.

Thank you for the comment. The concept of surrogate/meta-modeling has been indeed around for a long time. However, while one line of work relates to developing novel mathematical approaches to construct accurate metamodels, other relates to its use for purposes such as evolutionary search in order to reduce computational effort often referred as model management. The proposed work contributes towards the latter. Till date, majority of the works consider utilizing surrogate/metamodels to expedite the search for single-objective problems (SO), and there have also been increasing number of recent studies in dealing with multi-objective optimization problems (MO) up to 2/3 objectives. However, as far as many-objective problems (MaO) problems (> 4 objectives) are concerned, the foray of surrogate-assisted methods is recent and so far there are only two published works [1, 2] on it (the former is applicable to unconstrained problems and the latter to constrained problems where constraints are evaluated and considered cheap). The key difference in our proposed approach lies in the process of *selection schemes* (i.e. which designs to retain and promote). Surrogate assisted SO algorithms work (typically) based on ranking based on a singular measure (which can be predicted or evaluated), whereas MO problems typically work based on Pareto-dominance (of evaluated or predicted objective values). In our previous work [3], we had proposed and analyzed surrogate assisted optimization strategies for MO problems, comparing and contrasting them on various 2-3 objective problems. However, fundamentally, these strategies cannot

scale up to MaOPs because of non-dominated sorting involved which cannot induce sufficient selection pressure for MaOPs an aspect that is well reported in the literature [4]. This necessitates development of surrogate assisted MaOP approaches where the key lies in the ability to select-improve-evaluate limited number of solutions during the course of evaluation. The performance of the algorithm is directly affected by the choices and design of the above schemes which are detailed in the paper.

- When authors list their contributions in Section 2.3, out of 4 bullet points, item 3 and 4 are not contributions but rather benchmarking and application of the proposed approach to an engineering problem. Item 1 needs an extensive discussion to justify as a contribution (as I mentioned in the previous comment) and there remains Item 2 only. Authors need to rethink their contributions and revise their work appropriately.

  Thank you for the comment. We have now articulated the contributions with a clear focus on two, as we agree benchmarking and illustration can be construed as necessary to establish the claims. However, we would also like to highlight that to the best of our knowledge this is the first study that reports results of a surrogate assisted many objective optimization algorithm based on hypervolume metric in addition to inverted generational distance for limited number of function evaluations. This comparison has been now added in the revised draft based on comments from the reviewers which further supports our claims.

  It is important to take note that the first set of numerical benchmarks are all unconstrained problems (DTLZ and WFG). Our intent was to convey that the study [2] evaluated all constraints during its course of search i.e. considered constraints as computationally cheap. In our proposed algorithm, the constraints are considered to be computationally expensive and hence surrogates are constructed to represent each of them as done for the objectives. This comes into play for the engineering design optimization problem and we believe this is important for practical problems as many of the constraints could emerge from FEM or CFD analysis.

- Authors need to discuss several aspects of surrogate modeling in their paper. To make their work applicable to engineering design problems, they need to discuss how to determine an appropriate sample size under limited computation time. Also they need to discuss what is the required accuracy from these models and how the accuracy affect the solution quality. They probably need more results to demonstrate each.

  Thank you for this comment. Determination of an appropriate sample size for a given computational cost would be dependent on the non-linearity of the underlying objectives and constraint functions of the problem. Furthermore, when surrogates are used within an evolutionary framework, all choices related to population size, number of reference directions, number of solutions evaluated in every generation for model update and identification of solutions to improve/evaluate affect the performance of an approach. For the benchmarks, we have used the same set of parameters for a consistent and fair comparison with literature, while for the engineering design optimization problem we have presented the results of the approach with and without surrogates across a range of function evaluations. The results clearly illustrate the benefits of the surrogate assisted approach.
  While we have not activated "accuracy control" in this study, the algorithm does have a provision to "abandon approximations" and evaluate solutions if (a) the MSE of the best surrogate is greater than an user prescribed threshold or (b) a solution is beyond a user prescribed normalized distance from its closest evaluated point. These two schemes are designed to address the issue with respect to accuracy of representation that is necessary for real applications where there could only be enough computational resources to run the exercise once [5].

- In the application of the proposed method (Section 4.2.2), the authors do not give any insight into why SaMaO performs better than MaO in that problem; therefore it is not possible to understand in what cases a surrogate modeling approach is useful. As authors can also appreciate, the proposed problem is quite simplified form of real-world engineering problems and in practice it is important to know what happens when there is not enough samples or if the accuracy of the surrogate models is not sufficiently high.

  Thank you for this comment. In-fact, the intent of Fig.5 was to show that even at low function evaluations, the performance of SaMaO is better than MaO. We had shown the percentage of non-dominated solutions achieved by the two approaches (MaO & SaMaO) to demonstrate this. However, we understand that this may not be sufficient to clearly distinguish between the performance of the two algorithms. Therefore, in the revised manuscript, we have *additionally* included Hypervolume to show that the quality of set achieved by SaMaO is better than MaO right from the beginning. The reason is that given the same population size, both algorithms consume the same evaluations in the first generation, but thereafter SaMaO uses prediction based on the metamodels constructed through these solutions to predict better regions to sample. MaO on the other hand is driven purely by the fitness of the solution, which results in slower con-

vergence and consumption of higher number of evaluations at each generation. We do agree that the accuracy of the underlying surrogate would play a major role. To this end, having multiple types of surrogates in our proposed approach as opposed to only Kriging certainly offers a greater flexibility.

- Limitations must be stated and authors need to comment on the general applicability of the proposed method.

  Thank you for this very important comment which has allowed us to undertake a critical reflection. We have included the limitations of the work in the revised manuscript for completeness which sets the tone for future studies. Briefly, these include - (a) the approach does not include any specialized methods to handle large number of variables (b) no mechanisms to deal with missing data due to failed simulations and noise, an aspect that is increasingly coming into light and (c) more sophisticated techniques to handle highly irregular shapes. Some of these directions are currently being pursued by us.

Additional Comments:

- Vehicle architecture design does not accurately represent the illustration example. In general, product/system architecture design refers to a set of discrete choices that define how the building blocks of the system are connected and interact with each other. The illustrated example has only continuous design variables defining the size of the building blocks + cost. It is misleading for people who work in that area and I suggest the authors to use the phrase vehicle design by omitting the word architecture.
  Thank you for the comment. We have modified the terminology as suggested in the revised manuscript.

- The authors need to define regular and irregular PFs in the introduction. They promised to discuss later but there is no clear definition in the paper. It is necessary for a general audience.

  Thank you for the comment. We have added this information in the revised manuscript.

- The discussion on existing EA in Section 2.1 is not clear. The second paragraph in that section discusses what these methods do up to ranking, but since there is no discussion what they do after that, the sentence . therefore severely diminishing any selection pressure for the population to migrate towards the PF is hard to understand for a general audience.
  This is an important discussion that the authors allocate a large portion and clarity is important for readers to appreciate the value of MaOP.
  Thank you for the comment. We have revised the statement which now offers greater clarity.

- Decomposition in optimization is used in different meanings and the authors use it for a very specific meaning here. In Section 2.1, it is important to clearly define what is decomposed (i.e, what are these subproblems). For instance, they are clearly not decomposing the problem with respect to design variables (which is one meaning used in the optimization literature) but instead, decomposing the problem with respect to objective functions along reference axes. Authors need to clarify that.
  Thank you for the comment. We have clarified in the revised manuscript that decomposition refers to objective space in this work.

- In Section 2.1 Both of these distances are calculated in a normalized objective space using the reference point and the nadir point . It is not clear what is nadir point used for. My understand is that it is used to normalize the space but the sentence might also mean something related to the calculation of the reference vectors. it needs to be clarified.
  Your understanding is correct; the nadir point is used as the upper bound (1) of the normalized space. We have clarified this in the revised manuscript.

- Equation (9) should be for f5. There is typo there.
  Done.

- How is loc_seating calculated?

Reviewer 2

This paper considers an important problem of multiobjective design optimization and proposes a surrogate-assisted evolutionary optimization algorithm to arrive at a set of Pareto optimal solutions. While the paper is generally well written, I find that the original contributions of the paper are limited. The paper will be acceptable to me with major revisions once the following problems are addressed.

- I found the literature survey somewhat lopsided. On the one hand the authors do a great job of reviewing literature in evolutionary multiobjective design optimization, on the other, they do not consider the literature in the classical optimization or multifidelity area. While considering competing surrogate models, some comparisons of efficacy/efficiency needs to be made.

  Thank you for this observation. We have referred to a review paper [6] in the introduction, which covers the use of metamodeling in various methods, including classical optimization, in the context of engineering design. We have clarified this in the revised manuscript and also added citation to another paper [7] which discusses the comparative behavior of different metamodeling techniques. Given that the focus and intended contributions of the paper are towards an evolutionary approach, we have discussed them in more detail given the space limitations.

- The authors need to motivate the readers well as to why yet another evolutionary search procedure is needed. The authors also do not make a compelling case for evolutionary algorithms. If expensive function evaluations is the motivation, why not just go with surrogates and perform gradient based optimization?

  Thank you for the comment. The motivation behind using evolutionary search procedure is highlighted on page 1, which we have refined further to make clearer: "For solving generic MOPs, metaheuristics, e.g. evolutionary algorithms (EAs) are commonly used [8] as they do not suffer from limitations of classical techniques such as gradient-based solvers [6]. Being inherently population-based, EAs are well-suited to solving MOPs since the population can approximate the PF in a single run instead of point-to-point (including gradient-based) methods which need to be run repeatedly to cover the PF. Furthermore, they do not require mathematical properties such as continuity and differentiability as pre-requisites, and can be applied to highly non-linear or even *black-box* (where objective/constraint functions are hidden) optimization problems. For general applicability, we assume all objective/constraint functions to be black-box in this work, and therefore considering the above factors, EAs are, in principle, suitable methods to solve them".

  There are also some other concerns in using gradient based optimization, which we have not discussed due to space limitations. For example, gradient based optimization is inherently designed for single-objective problems and there is no direct/"best" way to extend it to multi-objective optimization. Scalarizing multiple functions into one could have undesirable effects, such as bias towards particular objective(s) and achievement of corner solutions instead of uniform distribution for non-convex Pareto fronts. Another way of using single objective search to solve multi-objective problems could be the concept of $\varepsilon-$constraint; where each objective is individually optimized by setting others are different constraint levels. This again involves choosing the threshold values for the constrained objectives, the range of which is unknown before the optimization. Moreover, for large number of objectives, correspondingly high number of single-objective problems need to be solved as well, which is also computationally expensive.

  It is however, worth noting that the concept of decomposition partially uses the idea of classical optimization, where the single-objective "sub-problems" along diverse reference vectors are solved in a cooperative manner. Additionally, in our proposed algorithm as well, local search using a (numerical) gradient-based method is performed in the "Identify" phase on Page 5. Thus, elements of classical optimization are indeed helpful even in the evolutionary approaches.

- While engineering design problems do involve many attributes, most of the times, they can be reduced to a smaller number. The authors need to highlight better the cost-benefit aspects of running an algorithm such as theirs and the challenge of implementing it.

  Thanks for the comment. Indeed, it is true that in some cases the objectives of the design may have strong correlation, due to which considering all of them may not be necessary, and the problem complexity (number of objectives) can be reduced without sacrificing the obtained results. However, this is not always the case, so it is still worth developing more generic algorithms that can handle large number of objectives. In one of our previous works [9], we have discussed this issue in detail and developed a method to come up with reduced objectives, where possible. It is worth noting that in some cases (e.g. the case of DTLZ1-3 benchmarks), even omitting one objective incorrectly would lead to the collapse of the remaining Pareto front (i.e. the multiobjective search will yield only one solution instead of a surface); whereas in some other cases, a smaller subset of objectives could be sufficient to obtain the Pareto front. Sometimes, even after

reduction, the dimensionality remains large (e.g. radar waveform optimization problem of 9-objective reduced to 6 in [9]). Since an accurate identification of dimensionality itself involves substantial number of function evaluations (in most of the existing works including ours), it is not an attractive choice for cases where the function evaluations are expensive.

That being said, if the number of objectives could be reduced simply based on prior knowledge or availability of analytical/empirical functions or correlation information for some of the objectives, then the problem formulation could be simplified to some extent during modeling itself. In our algorithm, in order to avoid bias towards the given problems, we have always assumed them to be *black-box*. In either case, a good algorithm will nonetheless be required to solve it, so the effort in developing it is still beneficial.

- Decision analysis and utility theory in general provide a normative way of multiattribute decision making. In fact, there are sophisticated techniques available in multiobjective utility assessment. The criticisms that the authors note of using a single combined objective (on page 1) do not stand this scrutiny. They need to better address this issue or at least acknowledge decision analytic methods.

  Thank you for this comment. As suggested, we have added reference to some of the related works; but since the focus in the paper is on improvement of Pareto front approximation rather than choice of solutions, we have not discussed them in detail for the sake of brevity.

- I was disappointed that the authors did not provide good performance metrics for computational effort required in each case (time per iteration etc.). While, the same number of functional evaluations was provided, what was the computational effort of running the algorithm itself?

  Thanks for the comment. We provide the justification for this in Section 1. Since the algorithm is aimed towards computationally expensive problems, we assume that majority of the time is taken up by the simulation (e.g. 400 CPU hours for a single scramjet shape flow analysis). Therefore, the time taken by the internal components of the algorithm, such as crossover, metamodeling, are considered negligible when compared with the time taken to evaluate a single simulation. The proposed approach is specifically designed for such classes of problems.

- The benchmarks problems chosen are the strong point of the paper.

  Thank you for the comment.

- The vehicle architecture problem, on the other hand, was not involved enough in my opinion. Five objectives, even according the Debs paper which the authors reference, are not that difficult for the types of objective functions used. At the same time, I think some of the objectives are going to be dependent on each other (cost, room, weight), making the problem easier than it first seems.

  Thank you for the comment. Recent studies and some of the reviews (e.g., see [4]) in multi/many-objective optimization demonstrate that the proportion of non-dominated solutions in a randomly initialized population is already significant enough to weaken the selection pressure in Pareto dominance based objectives (such as NSGA-II) beyond four objectives. We have therefore tested on a wide variety of benchmarks to ascertain the performance of the algorithm. With the engineering design problem, it is to be noted that all objectives and constraints are considered to be in *black-box* form i.e. information about gradients, linearity etc. are *not* known to the algorithm. We have presented the correlations in the Figures between the functions which clearly show that not all have high positive correlations.
  There are two other aspects that we would like to draw attention towards. The first is the fact that for unconstrained many objective optimization problems, existing nondominance based methods that are highly successful for 2-3 objective optimization problems may not deliver well converged solutions purely due to lack of convergence pressure. The unconstrained benchmarks discussed in the paper clearly belong to this category. As for the engineering design problem, even though it has 5 objectives, the presence of constraints could actually turn the problem easier to solve. The lack of convergence pressure in the unconstrained version can be overcome by a constraint violation pressure in the constrained form until all solutions in the population turn feasible. Once all the solutions in the population are feasible, nondominance will fail to induce any selection/convergence pressure. Since for engineering design problems, we expect active constraints at play, multiobjective algorithms(2-3) may still deliver converged solutions in only a part of the Pareto front as opposed to a well distributed set of converged solutions delivered by many objective approaches.

  Miscellaneous/Typos

- I found the large data tables somewhat distracting. They can be summarized in smaller tables.

We have moved the large tables and figures to the supplementary file and have presented the tables highlight results of the signifcance tests in the main text.

- I am sure Figure 6, in the middle of the references is an oversight. This should be fixed.

  [Check formatting options after revision].

Reviewer 3
Summary
This paper introduces a new surrogate-assisted evolutionary algorithm for many-objective optimization of expensive objective functions. The algorithm is based on the concept of decomposition and it has a number of advanced features such as adaptation of reference vectors, a pool of surrogate models, promotion of marginally infeasible solutions, etc. The performance (inverse generational distance) of three different versions of the algorithm using different combinations of metrics for selection of solutions is compared against state-of-the-art algorithms in a large computational experiment with a range of well-known benchmark problems. In addition, the algorithm is compared against a version of the same algorithm without the surrogate assistance component on a real-world vehicle architecture design problem.

Recommendation
This is overall a strong paper and I anticipate I will be recommending it for acceptance after a round of reviews. The paper addresses a relevant and important problem, namely the optimization of many conflicting objectives calculated through expensive objective functions (e.g., CFD, FEA). The background and literature review are adequate and exhaustive. The contributions are clearly articulated, and the experiments are well designed and conducted. The paper is well-structured and well-written.

However, I think the paper could be improved on a number of fronts. Please see the high-level and low- level comments below.

Thanks for your positive comments and constructive criticism.

High-level comments

- My primary criticism of this paper is that the number of function evaluations used in the real- world problem (4000) seems unrealistically high. The whole point of the paper is to propose an optimization algorithm for cases in which objective functions are very computationally expensive, such as FEA or CFD simulations, which could require hours of computational time per solution. Why not do this test with a more reasonable number of function evaluations, perhaps 400? The problem is that by not doing that, the paper leaves the reader wondering whether the method actually scales down well; perhaps many true function evaluations are needed to obtain accurate surrogates? The benchmark problems are done with a much smaller number of function evaluations, but the Pareto fronts of these synthetic problems are likely to be smoother and easier to interpolate than real simulations.

  Thanks for the comment. The number of "allowed" function evaluations can be subjective as it really depends on how expensive the problem at hand is. However, to demonstrate that the proposed algorithm scales down well, we have now included results for different stages - 500, 1000 and 4000 evaluations for the vehicle design problem. Due to space limitations, these results are included in the supplementary material. The results show that even for lower number of evaluations, the algorithm performance is better than the version without surrogate.

- Its also disappointing that no actual simulations were used in the real-world problem. Certainly, simulations are more computationally expensive than empirical models, but this is not the issue, since in the end the authors are considering a fixed number of function evaluations. The problem is that real simulations are nasty: they have noise (numerical errors or other), they fail (leading to missing data), etc. These factors likely make the task of the surrogate model more difficult, leading to an increase in the number of simulations needed for the fitting process.

  Thanks for the comment. Certainly, simulation-based optimization has several other challenges including noise and missing data, which are not dealt with in this work. However, computational expense is also one of the major concerns as highlighted by the literature in this area, and the prime focus of this work. The use of fixed evaluations (which can be very few) represents the constraint on time available for optimization; and in case of computationally expensive problems (or even physical-experiment based problems) this time is almost wholly consumed by the evaluation of

objective/constraints. While we appreciate than an example with simulations would have been more attractive, its often difficult to be used for future comparison and validation due to commercial/in-house proprietary analysis tools. Our intent is to draw attention of potential industrial users where they can see the link and almost certainly will have far more complex/nonlinear objectives and constraints involved in their design optimization process.

As for missing data and noise, there are indeed some works addressing them. For example, one of the ways to handle it could be to introduce effective penalties for the regions where the simulations fail, as suggested in [10]. We acknowledge the limitation of the proposed method, and consider it as further refinements in the future work.

- Why did the authors choose IGD only as performance metric for the first problem? Why not look at hypervolume as well? Do these two metrics correlate perfectly?

  Thanks for the comment. IGD and Hypervolume both try to capture an overall (combined) measure of convergence *and* diversity, and are therefore somewhat (but not perfectly) correlated. Also, IGD, although a very good and widely used measure, is not Pareto-compliant (i.e., a point that dominates doesn't necessarily have better IGD). Hypervolume, on the other hand is Pareto-compliant but computationally intensive for large sets and high number of objectives. For completeness, we've presented both in the revised manuscript following your suggestion.

- Similarly, why did they measure performance with the fraction of non-dominated solutions in the real-world problem instead of IGD, like with the benchmark problems? It seems like an arbitrary change.

  Thanks for the comment, point taken. In the revised version, we've shown the change in IGD and HV metrics over the generations.

- Continuing with the benchmarking theme, why not compare SaMAO to MaO in the benchmark problems as well? Why just do that in the real-world problem?

  Thank you for the comment. Since KRVEA already showed consistently better results than MaO in the original study [1], we omitted comparison with MaO for the benchmark problems to limit the volume of data in the tables.

- The authors seem to have taken some of the results (RVEA, ParEGO) from the literature. I know it was mentioned earlier that some of the parameters for the proposed algorithms were taken from the literature, but it would probably be good to add a sentence indicating very clearly that the results taken from other papers used parameters that are perfectly consistent with those used in this paper for the other methods, so that it is clear that we are comparing apples to apples.

  Thank you for the comment. We have clarified it in the revised manuscript.

- The authors claim that ASF-ED is "clearly" the best method based on the results of Table 5. However, unless I am reading the table incorrectly, ASF-ASF and ASF-ED seem to be almost equivalent in terms of face-to-face performance, with a slight advantage for ASF-ASF actually (ASF-ED is significantly better once, ASF-ASF is significantly better 4 times). The same argument could be made for Table 3. Although in the case of Table 3 there seems to be a slight advantage for ASF-ED, claiming that these results "clearly" or "convincingly" show that ASF-ED is better seems like a very strong claim not supported by results.

  Thank you for this observation. We have had a re-look into the results and modified the text to reflect this.

- While it is fairly clear what the MaO algorithm is that was used as control in the real-world problem, it would probably be good to clearly articulate the differences. Is it same pseudo-code as in Algorithm 1 but without updating surrogates and always evaluating instead of approximating, or are there any other changes?

  Yes, the algorithm is the same, with always evaluating instead of approximating. This has been clarified in the revised version.

- This is a paper that would fit perfectly in other venues such as IEEE Transactions on Evolutionary Computing. However, for this community, some people may expect to see a comparison between the proposed evolutionary optimization and a more classical approach such as gradient-based optimization, or the more recent Bayesian optimization. In all fairness, I do not think the authors should do that, I am just saying that some people in the community may be of that opinion.
  Our intent here is to bridge the gap between the Evolutionary Computing community and the Mechanical Design community so that the the benefits of development in the former can be realized by the latter. Secondly, the views from the

Mechanical design community can be used to drive developments within the EC community so that practically useful methods emerge with time. We are conscious of this and actively publish in both arenas and value feedback from both ends.

Low-level comments

- P3C1 "need further development" should be needs
- P4C1 "but is not" should be "but it's not"
- P4C2 The sentence "Subsequently (and whenever,... " appears to be incomplete
- P5C2 "it's neighboring reference directions" should be "its neighboring reference directions"
- There is no reference in the text to Figure 2. Is the ref to Fig 5 in page 7 actually referring to fig 2?
- P10C2 typo in "are run for run for a fixed budget"
- P11C2 "(-)ve values" ?? Why not "(-) values" or "negative values"?

Thank you for identifying the above mentioned errors in text/figure referencing. They have been corrected as per your suggestion.

- Why are population sizes different for different methods?

There was an error in this statement. The population sizes are the same for both SaMaO and MaO. The initial set sampled in both algorithms is also kept as $11 \times (n_{var} - 1) = 109$ as per the recommendation in [1]

- In "we observe the percentage of non-dominated solutions delivered by the two approaches at different stages during the search with respect to the final (combined) set of non-dominated solutions achieved", the final set contains the result from all 25 runs, right? Assuming you did 25 runs as you did for the benchmark problems. That is why the two fractions don't add up to 1 in Figure 5, correct?

For the vehicle design problem, the results are based on a single run. The final set contains combined set of non-dominated solutions obtained from the run of MaO and SaMaO. Therefore in the intermediate stages, the fractions do not add up to one, but at the end of the run, they add up to one.

Reviewer 4

General Comments:

In this paper, the authors present a surrogate assisted optimization algorithm for MOP/MaOP given a limited computing budget. Multiple types of surrogate models are applied and the one with the smallest MSE is chosen, and decomposition based EA algorithm is applied to solve the problem.

The details about this work are clear. However, some of the reviewer's concerns need to be addressed.

Major Comments:

- The contribution of this work is not significant. In Section 2.3, four contributions have been listed, however, claims 3 and 4 should not be considered as contributions. Claim 1 is not significant since most of the work has been done in the authors previous work, and only the objectives are replaced with surrogates since later in Section 3 it is said that constraints are considered to be computationally cheap and hence never approximated.

We wanted to convey that the only study in constrained MaoP [2] evaluated all constraints during its course of search i.e. considered constraints as computationally cheap. For the first set of unconstrained benchmarks this aspect does not come into play. However, in our proposed algorithm, the constraints are considered to be computationally expensive and hence surrogates are constructed to represent each of them as done for the objectives. This comes into play for the engineering design optimization problem and we believe this is important for practical problems as many of the constraints could emerge from FEM or CFD analysis.

With respect to the contributions, the vast majority of the works consider utilizing surrogate/metamodels to expedite the search for single-objective problems (SO), and there have also been increasing number of recent studies in dealing with multi-objective optimization problems (MO) up to 2/3 objectives. However, as far as many-objective prob-

lems (MaO) problems (> 4 objectives) are concerned, the foray of surrogate-assisted methods is recent and so far there are only two published works [1, 2] on it (the former which is only applicable for unconstrained problems, while the latter considers computationally cheap constraints which is often far from reality). Thus our intention was to illustrate the performance on unconstrained benchmarks and introduce an algorithm that can deal with problems where both objectives and constraints involve computationally expensive evaluations. The key contribution lies in the process of *selection schemes* (i.e. which designs to retain and promote). Surrogate assisted SO algorithms work (typically) based on ranking based on a singular measure (which can be predicted or evaluated), whereas MO problems typically work based on Pareto-dominance (of evaluated or predicted objective values). In our previous work [3], we had proposed and analyzed surrogate assisted optimization strategies for MO problems, comparing and contrasting them on various 2-3 objective problems. However, fundamentally, these strategies cannot scale up to MaOPs because of non-dominated sorting involved which cannot induce sufficient selection pressure for MaOPs an aspect that is well reported in the literature [4]. This necessitates development of surrogate assisted MaOP approaches where the key lies in the ability to select-improve-evaluate limited number of solutions during the course of evaluation. The performance of the algorithm is directly affected by the choices and design of the above schemes which are detailed in the paper.

- In Section 2.1, the theoretical issues of decomposition based EA are listed. Does this work deal with any one of the issues or just inherit these issues?

  They are some of the contemporary challenges in the field, which are not the focus of this work, but are included in the discussion for the sake of completeness. We have also made separate efforts towards addressing them. For example, in [11], we present an algorithm that uses two sets of reference vectors adaptively, one suitable for "regular" Pareto fronts and another for "inverted" Pareto fronts. In [12], we present an algorithm with adaptive reference vectors that can conform to different types of Pareto fronts - disconnected, degenerated, constrained, highly convex/concave etc. Both the above works also included some proposals to deal with the dominant resistant solutions (DRS). Evidently, including all of these aspects will not only be infeasible to present in the limited space, but also will take away the focus from the contribution of the current work. We have therefore scoped out this work and not used all the above components in the algorithm to stay focused on contribution towards handling computationally expensive problems.

- In Section 3, some literature has pointed out that NBI is not able to cover the full Pareto front, and some modified NBI is able to handle this issue. Do the authors take this issue into consideration?

  Related to the discussion above, we have developed some ways of overcoming this limitation for generic decomposition-based algorithms in [11, 12]. However, we have not used them in this work to avoid too many moving parts and keep the focus on contributions herein. We have added a paragraph on some of the limitations at the end of Section 4.

Minor Comments:

- In abstract, line 7, referred to asmany-objective should be referred to as many-objective.
- In Section 2.2, To the authors knowledge, the only has been [10] needs to be revised.
- The last sentence in Section 2.2, different type of should be different types of.
- The first sentence in Section 4.2.1 needs to be revised.
- Fig. 5 in the fourth paragraph of Section 4.2.1 should be Fig. 2.
- Point I is not marked in Fig. 2.
- The first paragraph after Eq. (10), only one run for is necessary.

  The above errors have been corrected.

References

[1] Chugh, T., Jin, Y., Meittinen, K., Hakanen, J., and Sindhya, K., 2016. "A surrogate-assisted reference vector guided evolutionary algorithm for computationally expensive many-objective optimization". *IEEE Transactions on Evolutionary Computation,* **PP**(99), pp. 1–1.

[2] Chugh, T., Sindhya, K., Miettinen, K., Hakanen, J., and Jin, Y., 2016. "On constraint handling in surrogate-assisted evolutionary many-objective optimization". In Parallel Problem Solving from Nature, pp. 214–224.

[3] Bhattacharjee, K. S., Singh, H. K., and Ray, T., 2016. "Multi-objective optimization with multiple spatially distributed surrogates". *Journal of Mechanical Design,* **138**(9), pp. 091401–091410.

[4] Ishibuchi, H., Tsukamoto, N., and Nojima, Y., 2008. "Evolutionary many-objective optimization: A short review". In

Evolutionary Computation, 2008. CEC 2008.(IEEE World Congress on Computational Intelligence). IEEE Congress on, IEEE, pp. 2419–2426.

[5] Isaacs, A., Ray, T., and Smith, W., 2009. "Multi-objective design optimisation using multiple adaptive spatially distributed surrogates". *International Journal of Product Development,* **9**(1-3), pp. 188–217.

[6] Wang, G. G., and Shan, S., 2007. "Review of metamodeling techniques in support of engineering design optimization". *Journal of Mechanical Design,* **129**(4), pp. 370–380.

[7] Jin, R., Chen, W., and Simpson, T., 2001. "Comparative studies of metamodelling techniques under multiple modelling criteria". *Structural and Multidisciplinary Optimization,* **23**(1), Dec, pp. 1–13.

[8] Deb, K., 2001. *Multi-objective optimization using evolutionary algorithms*, Vol. 16. John Wiley & Sons.

[9] Singh, H. K., Isaacs, A., and Ray, T., 2011. "A pareto corner search evolutionary algorithm and dimensionality reduction in many-objective optimization problems". *IEEE Transactions on Evolutionary Computation,* **15**(4), pp. 539–556.

[10] Forrester, A. I., Sóbester, A., and Keane, A. J., 2006. "Optimization with missing data". In Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, Vol. 462, The Royal Society, pp. 935–945.

[11] Bhattacharjee, K. S., Singh, H. K., and Ray, T., 2017. "A novel decomposition based evolutionary algorithm for engineering design optimization". *Journal of Mechanical Design,* **139**(4), p. 041403.

[12] Asafuddoula, M., Singh, H., and Ray, T., 2017. "An enhanced decomposition based evolutionary algorithm with adaptive reference vectors". *IEEE Transactions on Cybernetics,* **in press**.