Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/authorsrights

Author's personal copy

Computational Statistics and Data Analysis 71 (2014) 1178–1192

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Optimal designed experiments using a Pareto front search for focused preference of multiple objectives

Lu Lu [a], Christine M. Anderson-Cook [b,*], Dennis K.J. Lin [c]

[a] Department of Mathematics & Statistics, University of South Florida, United States
[b] Statistical Sciences Group, Los Alamos National Laboratory, United States
[c] Department of Statistics, Pennsylvania State University, United States

ARTICLE INFO

Article history:
Received 26 June 2012
Received in revised form 8 April 2013
Accepted 10 April 2013
Available online 19 April 2013

Keywords:
Multiple criteria optimization
Prioritization of criteria
Beta distribution
Focused Pareto front search
Computational efficiency

ABSTRACT

Finding a best designed experiment based on balancing several competing goodness measures of the design is becoming more important in many applications. The Pareto front approach allows the practitioner to understand trade-offs between alternatives and make more informed decisions. Efficient search for the front is a key to successful use and broad adoption of the method. A substantial computational improvement that conducts a more focused search when the experimenter has a focused *a priori* preference for the prioritizations of the multiple criteria is described. By utilizing a user-specified desirability function weight distribution for quantifying the preferences on different criteria, an algorithm to efficiently populate the desired portion of the front for two-criterion optimization is developed. Improvements over the full Pareto front search for completeness of the front in the region of interest, computational efficiency, and variation of the search are demonstrated with a screening design example where the objectives are precise model estimation and capability to protect against model mis-specification. Much of the existing literature focuses exclusively on finding the Pareto front, but does not offer strategies for making a choice of a best solution from the rich set of options identified on the front. A streamlined decision-making process with a set of tailored graphical tools to facilitate an informed and justifiable decision is described. The graphics incorporate *a priori* focused prioritization of the criteria, and hence are helpful to match decisions to design goals.
© 2013 Elsevier B.V. All rights reserved.

1. Introduction

In many design selection/construction situations, better overall designs can be found when diverse objectives are simultaneously considered. Two alternatives have been developed for this application: The desirability function (DF) approach (Derringer and Suich, 1980) typically searches for a best desirability score value based on a single subjective choice of fixed weights for quantifying the contributions from the different criteria specified. Lu et al. (2011); Lu, Anderson-Cook et al. (2012) describe an alternative process for finding a suite of best designs using the Pareto front (PF) approach. There are unique challenges for design selection compared with general multiple response optimization due to the discrete nature of many design criteria from a finite candidate set of design locations, the non-smooth change in design performance due to the inherent interdependence of design locations, and the existence of isomorphic designs (see a more detailed discussion in Lu et al. (2011)). The strategic search algorithm proposed in Lu et al. (2011) populates the PF to objectively eliminate inferior choices in the first phase, and then incorporates the experimenter's goals to select a best design after assessing performance

* Corresponding author. Tel.: +1 505 606 0347.
E-mail address: c-and-cook@lanl.gov (C.M. Anderson-Cook).

0167-9473/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.csda.2013.04.008

Author's personal copy

and robustness to different possible weightings. Using the PF allows the practitioner to explore trade-offs between different candidate solutions, and better understand what choices are possible.

The DF and PF approaches represent two decision-making extremes for the flexibility of choices versus computational efficiency. The DF search is fast and efficient since it looks at a single functional form with fixed weights for the criteria. There have been a few alternative functional forms for combining criteria proposed in the literature, such as in Harrington (1965) and Derringer and Suich (1980). The appropriate DF form should be chosen carefully based on the application-specific characteristics of the objectives. However, regardless of the functional form, the DF approach (1) requires the user to specify in advance the range of criteria values on which to scale, and which weights to use for combining criteria in the search; (2) provides a single answer which gives no frame of reference for calibrating the goodness of the solution found; and (3) does not allow for easy exploration of the robustness of the solution to different weights when user preferences are subject to uncertainty.

The general PF approach identifies solutions best for any possible criteria weight combinations. It allows the user to select the functional form and weights for combining the criteria after identifying the collection of superior designs. However, the cost for this flexibility is that the search for the complete PF can often be time-consuming, even though some solutions may be known to not be of interest ahead of time. In other words, the DF approach is computational efficient by being narrowly focused with the solution potentially sensitive to the subjective weighting choices, while the PF provides maximal flexibility at the expense of additional computational effort by finding all possible solutions.

This paper proposes an alternative search and decision-making approach for situations with more focused preference of weighting space, which balances computational efficiency while still allowing a range of weights to be considered. This provides a timely solution while avoiding the risks of making an over-simplified decision by focusing on just a single subjective weighting choice. Our experience with design of experiment applications is that it is typically hard for a subject matter expert (SME) or a decision maker (DM) to specify a single set of weights for optimization, but they are more comfortable with specifying a range or distribution of weights for describing their preference for the criteria after an initial calibration of what range of criteria values are likely. When a user-specified range or weight distribution is available, explicitly searching for only solutions based on these weight combinations within the relevant region can streamline the search, while still allowing flexible exploration of alternatives and their trade-offs. The time-savings for the more focused search can be substantial, which can make considering the PF approach more appealing.

Approaches for finding partial relevant regions of the PF have been proposed in the literature. Wagner and Trautmann (2010) approximate a region of the front of interest using DF-based approach with some forms of Harrington's desirability function. This and the rich literature on PFs focus on algorithm development for efficiently finding the relevant region of the PF, but offer little advice about choosing a single justifiable final solution from the front. Lu et al. (2011) proposed a process with two distinct stages for improved decision-making considering the entire design space. First, the search for the PF is the objective stage for identifying and understanding possible choices. The second stage involves a process for making a defensible choice of a single best solution based on quantitative evaluation of alternative choices and understanding trade-offs. In this paper, we describe a new search algorithm for efficiently finding the focused region of the front, as well as how to adapt the second subjective stage for incorporating this preference in decision-making. The graphical summaries proposed in Lu et al. (2011); Lu and Anderson-Cook (2012) are modified and a new adaptation of the Fraction of Weight Space (FWS) plot (Lu, Chapman et al., in press) is developed to incorporate user-specified priorities of the experiment into a simple quantitative summary of individual design performance over the focused weighting space of interest. This enables better decision-making with easier comparisons of alternatives and is helpful for ranking the choices. In the next section, we provide some background before introducing the focused PF approach.

2. The general Pareto front approach

Pareto optimization for multiple responses has been extensively used (Coello Coello et al., 2007; Deb, 2001; Kasprzak and Lewis, 2001) with applications in many different fields (see Gronwald et al. (2008) and Trautmann and Mehnen (2009) for specialized case studies). The PF approach was introduced to the design of experiments paradigm by Lu et al. (2011). The approach is comprised of two distinct stages. The first stage finds the collection of not inferior Pareto optimal designs. The search for the PF is objective without considering any SME or DM opinion about the prioritization of the objectives, how to convert the criteria values to the DF scale, or how to integrate the criteria into an overall summary for ranking design performance. All contending designs are compiled with a Pareto search algorithm, and then used in the second stage to guide making a justifiable choice based on matching the design performance with experimental goals. Since a complete PF includes all contending designs across all possible weights, the experimenter has a better understanding of the interrelationship of the criteria, and can better choose weighting and scaling schemes for the criteria.

A key to the success of the general PF approach is to efficiently populate the complete PF. A traditional point exchange algorithm based on a candidate design point list uses multiple starts, and looks for a single solution from each random start by iteratively swapping out individual points with alternatives until no improvement can be made to all criteria simultaneously. The Pareto Aggregating Point Exchange (PAPE) algorithm (in Lu et al. (2011)) improves computational efficiency by evaluating all designs and accumulating the PF throughout the search process. Lu and Anderson-Cook (2012) refine the PAPE algorithm with a less greedy search updating mechanism and diversifying the search directions to give better coverage and reduced chances of getting stuck at local optima. Searching in multiple diverse directions using a grid

Author's personal copy

Table 1
Criteria values for the nine points on the complete PF based on *D*-efficiency and tr($\mathbf{AA}'$).

Design	(*D*-eff., tr($\mathbf{AA}'$))	Design	(*D*-eff., tr($\mathbf{AA}'$))	Design	(*D*-eff., tr($\mathbf{AA}'$))
1	(0.771, 2.345)	4	(0.886, 2.480)	7	(0.916, 2.944)
2	(0.797, 2.367)	5	(0.902, 2.618)	8	(0.928, 3.000)
3	(0.863, 2.420)	6	(0.907, 2.920)	9	(0.939, 3.333)

of weight combinations more quickly populates the PF. Note that all of the above search algorithms evaluate all individual designs for inclusion on the PF using their original criteria values. The DF metric used in the updating step is to direct the parallel searches to reach different regions of the PF more quickly and directly, and hence improve search efficiency. The completeness of the identified front is related to, but not dependent on the particular DF form used during the search. Based on our experience, when a fine grid of weight combinations are used for guiding the search directions, the result is quite robust to the DF form choice. More details of these algorithms and their comparison are available in Lu and Anderson-Cook (2012).

After building the PF, the Pareto decision-making analysis is conducted using graphical summaries. Consider one of the examples adapted from Lu et al. (2011) with the goal of finding a 14-run screening design for 5 factors with the goals of better precision of model estimation and protection from model mis-specification. The primary model of interest includes all main effects (*A*–*E*) and a subset of the two-factor interactions (*AB*, *AC*, *BD*, and *CE*). Although prior scientific knowledge suggests that these are the most likely terms, there is interest in protecting against one or more of the remaining two-factor interactions (*AD*, *AE*, *BC*, *BE*, *CD*, *DE*) being active. Based on this model, suitable candidate design locations are all combinations of the two factor levels $(-1, +1)$. Two criteria, maximizing *D*-efficiency and minimizing tr($\mathbf{AA}'$) are chosen to quantify the chosen design characteristics. *D*-efficiency, defined as $|\mathbf{X}_1'\mathbf{X}_1|^{1/p}/N$ where $\mathbf{X}_1$ is the design matrix expanded to the form of the specified model with p terms and N observations, is a common criterion for quantifying the precision of the coefficient estimates for a given model. For a specified model matrix, $\mathbf{X}_1$, with potential missing terms listed in $\mathbf{X}_2$ (see Myers et al. (2009, p. 284)), we define the alias matrix, $\mathbf{A} = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2$. Then, tr($\mathbf{AA}'$) measures the potential bias of the estimated coefficient if the specified model is inadequate. Smaller tr($\mathbf{AA}'$) values suggest designs with better protection against model mis-specification.

The complete PF based on the two criteria is found by using the PAPE algorithm and shown in Table 1 and Fig. 1(a). The front contains nine points, labeled designs 1–9, from worst to best *D*-efficiency. The Utopia point, defined to simultaneously have the best observed values for both criteria (maximum *D*-efficiency and minimum tr($\mathbf{AA}'$)), is shown in the bottom right corner of the plot. We see some disproportionate gains in one criterion by sacrificing from the optimum of the other criterion. For example, design 1 has the smallest tr($\mathbf{AA}'$) value. A slight sacrifice in tr($\mathbf{AA}'$) from 2.345 to 2.420 (about 3%) gives an alternative, design 3, with about a 12% gain in *D*-efficiency from 0.771 to 0.863. Hence, it is possible to achieve substantial gains in one criterion without too much sacrifice from the other.

Fig. 1(b) shows the best designs for all additive weight combinations of the two criteria using DF $= w_1c_1 + w_2c_2$, with $\sum w_i = 1$ and for c_i between 0 and 1 (scaled for the worst and best values of that criterion on the PF, respectively). Designs 1, 3, 4, 5, 8, and 9 are optimal for at least some sets of weights with increasing emphasis on *D*-efficiency. Designs 2, 6, and 7 are not optimal for any weights since they are not on the convex hull of the PF in Fig. 1(a). If another DF form is suggested by SME knowledge, then this plot can be easily adapted to reflect this form. From Fig. 1(b), when *D*-efficiency and tr($\mathbf{AA}'$) are equally important ($w_1 = w_2 = 0.5$), design 4 is best with considerable robustness around this specified weighting. As *D*-efficiency is given more weight, design 5 or 8 becomes the best. If *D*-efficiency is weighted more than 83%, then the best choice is the *D*-optimal design (design 9). Given Fig. 1(b), the experimenter can select a best design to match the goals of the study.

To better understand the performance of a particular design, Lu and Anderson-Cook (2012) developed a synthesized efficiency plot to show design performance relative to the best available for different weightings of criteria. The synthesized efficiency for a design is defined as the ratio of its DF score to the optimal DF value for a particular weight combination. Fig. 1(c) shows the synthesized efficiency plot for the six designs from Fig. 1(b). The 20 shades of white–gray–black summarize the high (white is 95%–100% efficient) to low (black is 0%–5% efficient) synthesized efficiency for different weights for an individual design. Designs 1 and 9 are both optimal for a single criterion, and hence have a white region with highest efficiency for that optimal criterion and a dark region where more emphasis is placed on the other criterion. Designs 3–5 all have a large region of white with no extreme dark shading (efficiency below 50%), which indicates well balanced performance across the entire weighting space. The synthesized efficiency plots help visualize the range and relative quality of good performance for individual designs.

The *fraction of weighting space* (FWS) plot (Lu, Chapman et al., in press) provides an overall quantitative summary of individual design performance across the entire weighting space and allows for easy comparison between competing choices. Fig. 2 shows the FWS plot for the six designs in Fig. 1(b) and (c). The line for each design shows for what fraction of the weighting space it has synthesized efficiency at least as large as a certain percentage of the best possible design performance for a particular weight combination. The calculation for the FWS plot is built from the set of synthesized efficiency values for a fine set of weight combinations in Fig. 1(c). The distinct efficiency values are found and sorted in descending order, and then plotted versus the fraction of weight combinations with efficiency at least as large as that value. The smoothness of the FWS plot is improved by using a finer set of weight combinations.

Author's personal copy

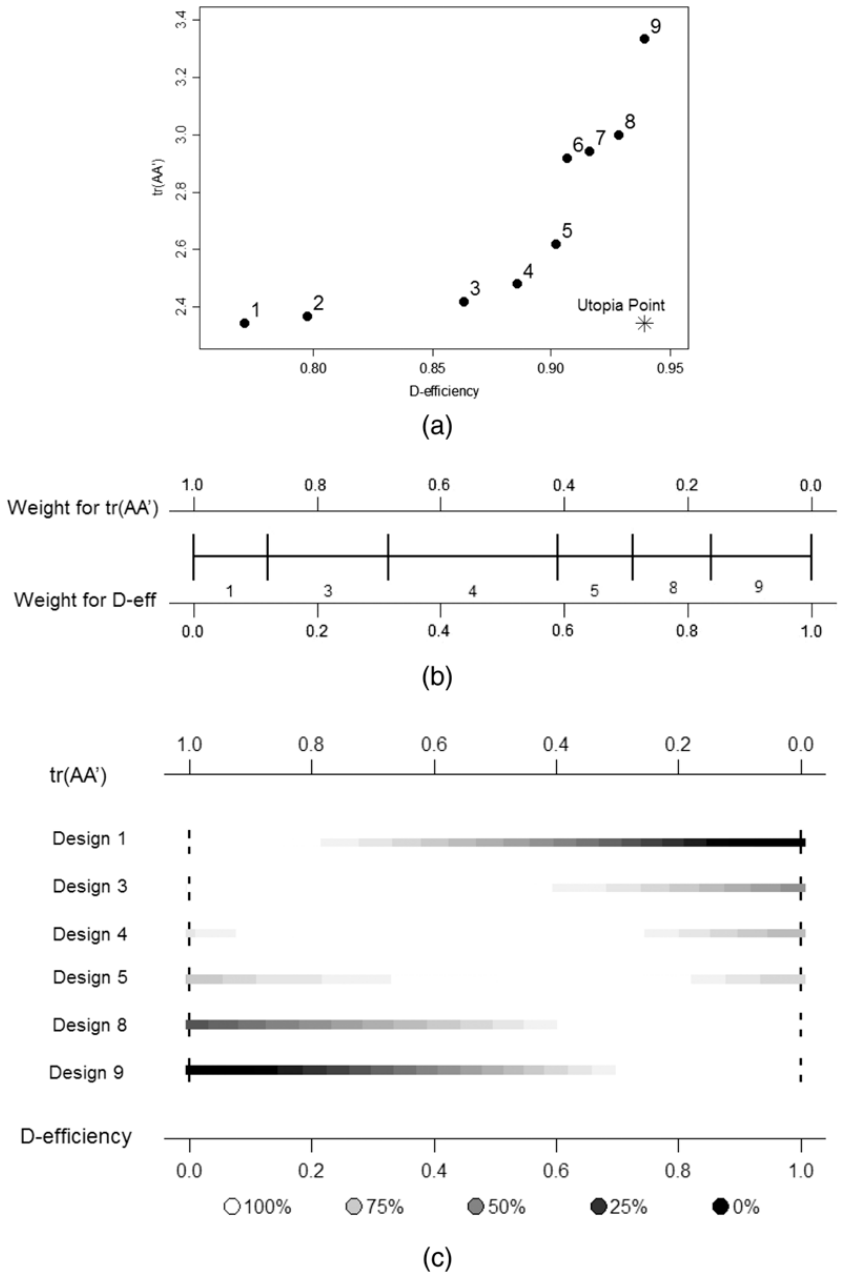


Fig. 1. (a) The complete PF based on *D*-efficiency and tr(**AA**′); (b) best designs for different weightings of the two criteria when additive DF form is used; (c) synthesized efficiency for different weightings for selected designs. The white–gray–black shading shows high to low synthesized efficiency.

From Fig. 2, we see that design 4 is at least 80% efficient for close to 80% of the weighting space. Design 5 has the best (highest) worst efficiency across the entire weighting space, but is not as good as design 4 for the majority of the weighting area. Design 3 has higher synthesized efficiency than design 5 for about 60% of the weighting space but has worse efficiency for the remaining area. Designs 1 and 9 each have efficiencies below 60% for about half of the weighting space. Design 8 has consistently higher efficiency than designs 1 and 9 across all possible weights but is uniformly outperformed by designs 3–5 for all fractions. The FWS plot adapts easily for any number of criteria and hence provides an easy dimension-free comparison between designs.

The above summaries assume no *a priori* information about whether some criteria are considered more important than others, and the set of weight combinations are chosen by evenly spreading them across the weighting space. However, there are situations when only a limited range of weightings are considered relevant and/or the weight combinations within the range are not equally likely. For example, consider that the experimenter wishes to emphasize precise model estimation with a weighting of *D*-efficiency between 60% and 80%, where the most likely weight is around 70%. Then a Beta (71, 31) distribution with approximate mean of 0.70 and standard deviation (sd) of 0.045 could describe the user's preference for weights and is shown as the double dashed line in Fig. 3. We use the parameterization where the mean and variance of a Beta(α, β) are $\alpha/(\alpha+\beta)$ and $\alpha\beta/((\alpha+\beta)^2(\alpha+\beta+1))$, respectively. The Beta distributions is a natural choice for quantifying the weighting preferences since it is familiar to many practitioners, and has a range of values between 0 and 1

Author's personal copy

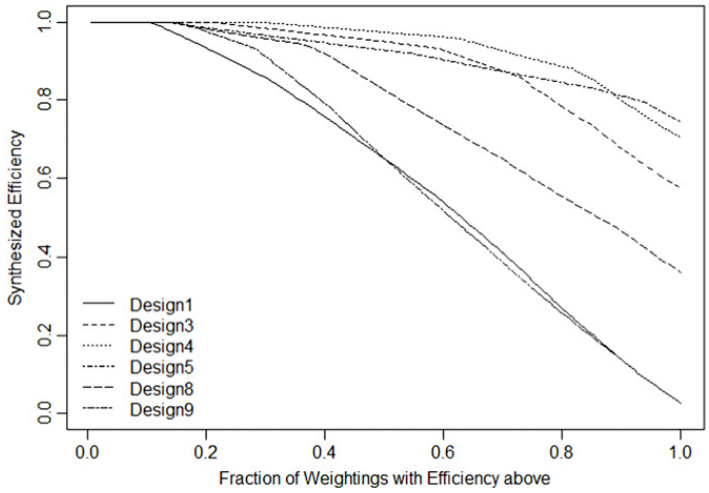


Fig. 2. The FWS plot for designs displayed in Fig. 1(c). The horizontal axis represents the fraction of weighting space that has synthesized efficiency at least as good as the corresponding efficiency value shown on the vertical axis.

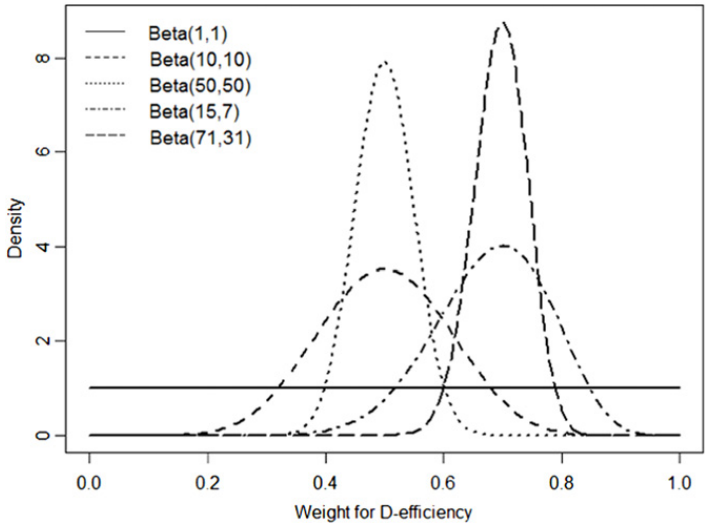


Fig. 3. Five different choices of weight distributions for *D*-efficiency.

with substantial flexibility in the shape, spread and symmetry of the distribution with different choices of parameters. For this example, the Beta distribution captured the SME's weighting preferences, but any other distribution could be used for other applications.

Other choices of weighting distribution are shown to illustrate some candidates for different users' preferences. For example, a Beta (10, 10) with mean $= 0.50$ and sd $= 0.109$ gives a wider spread of weightings for the two criteria, while a Beta (50, 50) with mean $= 0.50$ and sd $= 0.050$ has a more limited range. Compared to Beta (71, 31), Beta (15, 7) gives more emphasis on *D*-efficiency and considers a wider range of weights.

When only a limited region of weighting is relevant, seeking the complete PF search is not the most efficient use of computing resources. Hence, we develop a new tailored search algorithm which reduces search times by adapting the complete front search to find only the collection of solutions of interest. Details of the new search algorithm are given in Section 3. The *target front growth summary* quantitatively assesses search performance relative to the original complete front search in Section 4. In addition, graphical summaries are adapted to incorporate the users' potential non-uniform interest in the weighting space. These methods are illustrated in Section 5 using the screening design example. The impact on the search and obtained results for different user-specified weight distributions are explored and evaluated. Another possibility for improving the search algorithm is with stratified random sampling of weights (Lu and Anderson-Cook, 2012) to guide the search in diverse directions while allowing some of the uncertainty of weight specification to be included. Note that in this paper we choose to use Beta distributions for capturing the weighting preference of the two criteria due to its flexible shape of distribution, easy interpretation of parameterization as well as its natural scaling in the [0, 1] range. However, the focused Pareto search method and its advantage in computational efficiency are not restricted to this specific distribution. In addition, the method can be easily adapted to applications with different design scenarios and other criteria of interest. Section 6 has some discussion about adapting to related scenarios such as the use of multiplicative DF form and the stratified random weights for directing the search.

Author's personal copy

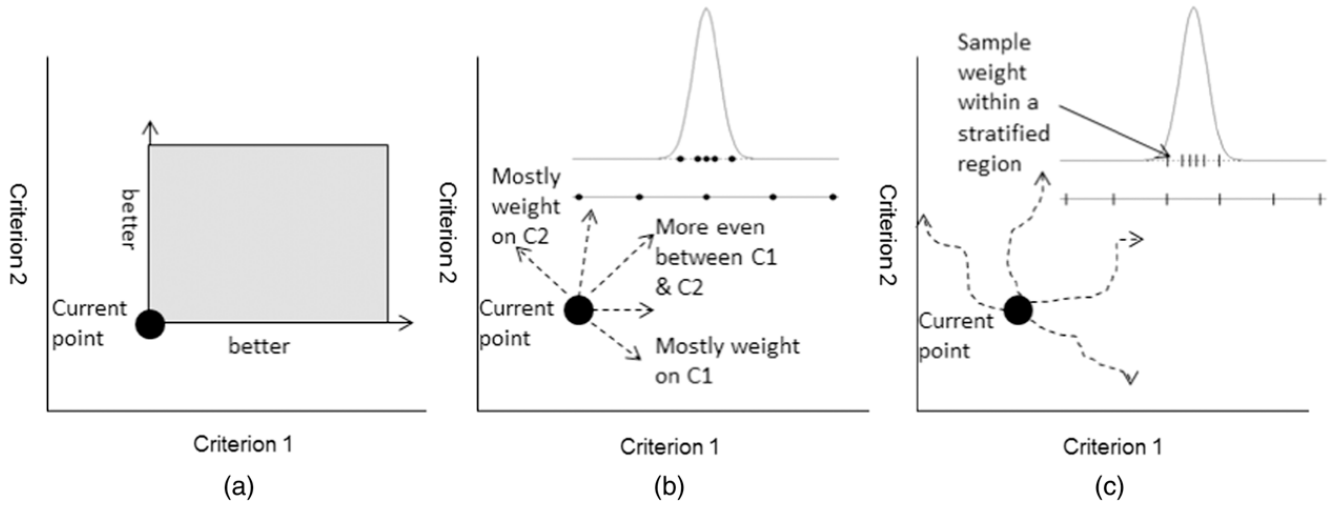


Fig. 4. Alternate updating mechanisms for two-criterion maximization scenario. (a) The greedy search approach can move only to an upper right quarter of the space. (b) The fixed weights approach can move to directions determined by all the weight combinations identified on the (0, 1) line segment. The upper right corner shows the weight combinations specified by a Beta (50, 50) and a uniform/Beta (1, 1) distribution by using 2.5%, 25%, 50%, 75%, and 97.5% quantiles. (c) The stratified random weights approach uses a stratified random sample of weights to determine the moving directions. The upper right corner shows the stratified regions partitioned by the 2.5%, 20%, 40%, 60%, 80% and 97.5% quantiles for a Beta (50, 50) and a uniform distribution.

3. Focused Pareto front search

This section gives details of the new focused PF search algorithm. The computational gains in its performance for finding the best solutions in the target range of the weighting distribution are demonstrated in Section 4 with the screening design example.

The original PAPE algorithm updates the current design (hence directing the search) by iteratively exchanging locations in the current design matrix with alternatives from a candidate set. Updating is based on finding strict improvement for at least one of the criteria without diminishing other criteria performance. We consider this to be a greedy search mechanism since it requires rigorous improvements based on all criteria simultaneously. This makes it difficult to update the current design and constrains the search to only move in a portion ($\frac{1}{2^k}$ for k criteria or a quarter for two criteria as illustrated by the gray region in Fig. 4(a) when the goal is maximizing both criteria) of all possible directions relative to the current design (denoted by the black solid circle in Fig. 4(a)). More random starting designs can mitigate the chance of getting trapped at local optima, but search performance is dependent on where the starting design falls in the design space. This leads to large variability based on the random starts. Lu and Anderson-Cook (2012) demonstrated that the greedy search method is an improvement over the traditional point exchange, but it can suffer from poor front coverage, low efficiency and large variability across multiple random starts compared to its enhancements.

One way of enhancing the search (suggested by Lu and Anderson-Cook (2012)) is to direct the search by using desirability functions with a fixed set of weights for the criteria. For each weight combination from the pre-determined set, a desirability index is calculated for every searched design to determine if the current design should be replaced by a new one. This directs the search toward designs best for that particular weight combination. Throughout the search, all designs are evaluated for inclusion on the PF based on their original criteria values. Consider the example with the PF shown in Fig. 1(a), a weight combination (1, 0) guides the search toward the D-optimal portion of the front, while a weight combination (1/2, 1/2) directs the search to the bottom right corner of the front. Designs best for slightly different weights than those targeted by the grid of weights are likely to be found as the search explores the space.

For the scenario with no prior preferences between criteria, the weighting space (represented by the line interval in Fig. 1(b) for the two-criterion case) can be explored by selecting weight combinations evenly spread over the entire space. For each random starting design, independent searches are conducted for all specified weights in multiple directions. The overall PF is updated based on all concurrent searches. By pursuing diverse, appropriately spaced directions, the search covers all regions, but also has minimal overlap of search paths for each random start. Also with the radiating searching paths, the performance is less dependent on the random starting designs. Therefore, the method has potential to find the complete front more efficiently and with less variability.

Note that the desirability function index used at each updating step is dependent on the scaling of the criteria values, which in turn depends on understanding the range of the criteria values. The best value for each criterion is easy to find with single criterion optimization. There are more possibilities for specifying the worst value (which maps to 0 in the scaled desirability). One simple choice is to use the natural worst value based on the definition, such as using the value 0 for D-efficiency for non-estimable designs. Another choice is to conduct some preliminary searches to optimize several subsets of the criteria, and then choose the worst value found from these results. For the two-criterion case, due to the unique ordering of points on the front, improving one criterion necessarily leads to a worse value for the other. Hence a

Author's personal copy

good strategy is to conduct preliminary searches for each criterion individually and scale the worst value found to zero for each criterion. However, this approach does not extend to cases with three or more criteria, and hence some alternative preliminary searches may be needed to gain some insights about sensible ranges of the criteria values.

Fig. 4(b) illustrates the general fixed weight updating mechanism. The solid black circle represents the current design at any step. The radiating arrows from the current design show multiple concurrent search directions determined by different weight combinations. The set of weights for the complete PF search is determined by evenly spreading the total number of weight combinations across the entire space. For a more focused Pareto search, we propose to use evenly spaced percentiles from the user-specified distribution.

The top of Fig. 4(b) shows possible weight combinations from a Beta (50, 50) distribution for quantifying the preference for D-efficiency (the weight for tr($\mathbf{AA}'$) is one minus the D-efficiency weight). The experimenter decides that the middle 95% of this distribution is of interest, and hence the weight for the D-efficiency criterion is bounded by the 2.5 and 97.5 percentiles. Within the range of interest, a set of evenly spaced quantiles specify the weight combinations from the chosen distribution, to divide the weighting space based on the probability distribution, and hence reflects a non-uniform preference on the weight combinations. The five black solid circles denote five weight combinations corresponding to 2.5%, 25%, 50%, 75%, and 97.5% quantiles from the Beta (50, 50) distribution.

The second horizontal line in Fig. 4(b) shows the weight combinations based on the same quantiles from a uniform weight distribution. This corresponds to the scenario with no *a priori* preference on the criteria but with less interest in the extremes. The complete search with no prior preference can be considered as a special case with a uniform weight distribution and extreme tail quantiles for specifying the range. Compared to the uniform distribution, the more concentrated Beta (50, 50) distribution puts the majority of search effort on the center of the weighting space. The number of quantiles determines the spacing between search directions. As the number of quantiles selected increases, the directions explored become more dense, but require additional time for each random start.

The focused fixed weight Pareto search process is summarized as below:

Step 1: *Specify the scaling scheme for converting the criteria values to the range [0, 1] and a function for combining the criteria into a single desirability function index* (typically, additive or multiplicative). Some preliminary searches based on each of the individual criteria can provide an understanding of the ranges of criteria values, or the user can use a natural worst case. Note that our experience shows that the search process is somewhat robust to the scaling choice within a sensible range, since the use of a fine set of weight combinations with good spread can compensate for the sensitivity of search results to the scaling scheme for an individual weighting choice.

Step 2: *Specify the weight distribution* (to describe the experimenter's *a priori* preference of different criteria) *and percentiles from the specified distribution* (for appropriate coverage of the focus region).

Step 3: *Randomly generate a non-singular starting design* of appropriate size from the candidate set of design locations using sampling with replacement. (For other problems which might require special design restrictions on candidate locations or design structure, special adaptations may be required for generating the starting design and the updating process.)

Step 4: *For each random start, conduct concurrent searches to improve the desirability score based on all specified weight combinations.* For each search, the current design is updated by exchanging rows in the design matrix with an alternative from the candidate set when the new design represents an improvement in the calculated desirability function index. More specifically, the criteria values of the current and new designs are scaled to [0, 1] based on the chosen scaling and their corresponding desirability function index values are calculated based on the specific weight combination. If the new design has a larger value than the current index value, then the current design is replaced by the new design. Each iteration updates the current design until no improvements are possible. A separate process evaluates all designs and combines all concurrent search results to populate the PF.

Step 5: *Repeat Steps* 3 *and* 4 *for multiple random starting designs.* The PF is updated to add new solutions as they are found and to remove inferior choices as new solutions improve the front. The updated PF is summarized after each random start.

An alternative to the fixed weight search is to use stratified random weights generated from the user-specified weight distribution. As shown in Fig. 4(c), the weight distribution is divided into multiple strata/regions with approximately even probability for each stratum. This can be done with a set of evenly spread quantiles from the distribution. At each updating step, a weight combination is randomly selected from each stratum. For each random starting design, multiple concurrent searches are conducted based on the random weight combinations generated from the weight regions. The stratification ensures diverse directions, but the random selection of a weight within each stratum allows small variation in direction. Like the number of quantiles chosen for the fixed weight search, the number of strata determines the spacing of the search directions. Lu and Anderson-Cook (2012) compare computational efficiency between fixed weights and stratified weights for the complete PF search. In this paper, we demonstrate the computational gains obtained from the focused Pareto search with fixed weights.

4. Performance of the focused Pareto front search

This section compares the focused and complete PF searches using fixed weights. The performance of the search methods are assessed using a PF growth plot (Lu and Anderson-Cook, 2012), which quantitatively summarizes three important

Author's personal copy

Table 2
Summary of times (measured in minutes) for finding the target designs on PF based on *D*-efficiency and tr(**AA**′) for different user-specified DF weight distributions.

Weight distribution	Target designs	5 quantiles/weight comb.		11 quantiles/weight comb.	
		Median	95% range	Median	95% range
Uniform/Beta (1,1)	1, 3, 4, 5, 8, 9	2.70	(0.74, 9.52)	2.60	(0.68, 7.47)
Beta (10,10)	3, 4, 5	0.93	(0.21, 2.60)	0.92	(0.18, 3.85)
Beta (50,50)	4, 5	0.49	(0.18, 1.21)	0.40	(0.14, 0.87)
Beta (15,7)	4, 5, 8, 9	0.35	(0.14, 0.87)	0.34	(0.16, 0.87)
Beta (71,31)	5, 8	0.34	(0.16, 1.04)	0.28	(0.17, 1.20)

characteristics of the search process: the completeness of the generated front, how quickly the search finds the front, and the growth variation associated with the use of multiple random starting designs.

We focus on the designs in the region of interest for evaluating search performance. Designs found on the front but not optimal for any weight combinations in the range of interest have been determined less relevant for the final decision, and hence are not considered. For the screening design example summarized in Fig. 1, the six designs shown in Fig. 1(b) are used to assess the performance of the complete PF search. For any focused non-uniform weight distribution, a subset of the six designs in Fig. 1(b) which fall within the bounding range of interest are considered as the desired target set. The second column in Table 2 shows the target designs for the distributions in Fig. 3. For example, a Beta (10, 10) distribution aims to find designs 3–5 that are optimal in the central region of the weighting space. However, when the narrower Beta (50, 50) distribution is specified, fewer designs (4 and 5) are important. The last two rows in Table 2 correspond to using the skewed distributions with more emphasis on the *D*-efficiency criterion (Beta (15, 7) seeks designs 4, 5, 8 and 9, while Beta (71, 31) seeks only designs 5 and 8).

To access the completeness of the identified front by a certain search algorithm, the best available front is needed as the "gold standard" for comparison. This should be obtained by combining information from all available sources for each particular application. Note that for a practitioner finding an optimal design for a specific problem, this stage of quantifying performance is not needed, but we include it here to illustrate the improvements possible. The efficiency of the search is evaluated based on the computer time needed to find the front. The associated process variation from the use of random starting designs is quantified by using a resampling method by sampling without replacement from a set of independent searches based on a small number of random starts (see Lu and Anderson-Cook (2012) for more details). The process for creating a front growth plot for a computing time-based summary of the proportion of the target designs found by a search method and its associated uncertainty is briefly described below.

First, N_b batches of searches, each with N_r random starts, are conducted using parallel computing. A PF is found for each batch of N_r random starts and the computing time used is recorded. An overall PF (typically used as a "gold standard" if no other information is available) can be obtained by combining all results from N_b batches.

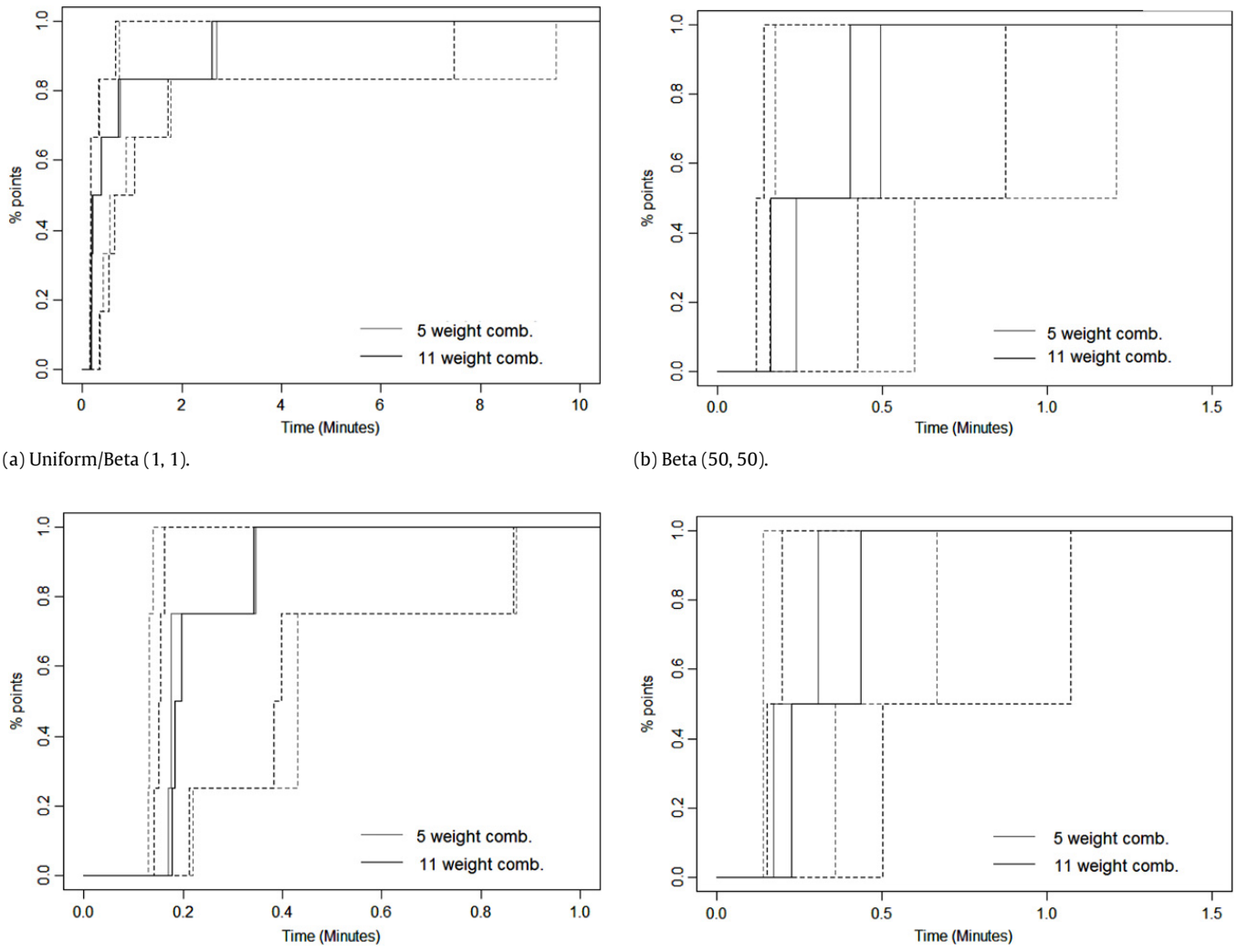
In the second step, N_s simple random samples with N_g batches in each sample are drawn from the N_b batches using sampling without replacement (since replicated batches do not contribute to the growth of the front). Note that N_g should be considerably smaller than N_b and a sampling fraction (N_g/N_b) between 10% and 20% is typically recommended. For each sample with N_g independent fronts, a new set of N_g fronts are obtained by sequentially accumulating the fronts. For example, the first new front is based on the first of the N_g batches. The second new front is the combined front for the first and second batches together. The proportion of the target designs found and the computational time are calculated for the N_g sequentially growing fronts. Note that the proportion is non-decreasing and computing time increases as more random starts are run. This progression of computer time and associated accumulated fronts are calculated for each of the N_s samples.

In the third step, the N_g proportion values are plotted versus computing times for all N_s samples using a step-function. The origin corresponds to having no front after no searches. The proportion jumps to a non-zero value with the first batch of results when at least one design in the target set has been found, and then remains flat until the time when at least one additional design in the target set is found. Different samples grow at different rates, which reflect the variation from random starting designs.

The last step summarizes the distribution of N_s paths with the median, lower (e.g. 5%) and upper (e.g. 95%) empirical quantiles of the proportion of designs found at any time. More specifically, all time points where any path has a jump are sorted in ascending order. At each of these times, the median number of designs found across all the paths is found. By connecting the set of all median values for all times, the median line across the time range explored is found. The lower and upper empirical quantile summaries can be similarly obtained.

Fig. 5 shows the front growth plot for the screening design example for different weight distributions for *D*-efficiency and tr(**AA**′). The scaling for both criteria was chosen so that the desirability scaled zero for both criteria was specified as the worst value on the PF. The results are obtained by using $N_b = 500$, $N_r = 2$, $N_s = 100$, and $N_g = 80$ for the uniform distribution and $N_g = 20$ for the Beta distributions shown in Fig. 3. The number of random starts within each batch, N_r, was determined by the relative speed to populate the front based on some preliminary runs. Typically a small number for N_r is used to smoothly display the progressive front growth process. The number of groups drawn in each sample was determined based on the time to find the all target designs for the worst case scenario summary.

Author's personal copy



(a) Uniform/Beta (1, 1).

(b) Beta (50, 50).

(c) Beta (15, 7).

(d) Beta (71, 31).

Fig. 5. PF growth plots for different weight distributions for *D*-efficiency and tr(**AA**′). The black and light gray lines represent searches using 5 and 11 fixed weight combinations. For each search, the three lines correspond to the 95% (dashed), 50% (solid) and 5% (dashed) empirical percentiles.

Fig. 5(a) shows the PF growth plot for a uniform or Beta (1, 1) weight distribution, where we seek to find all six designs in Fig. 1(b). The black lines correspond to the search based on 5 weight combinations (2.5%, 25%, 50%, 75%, and 97.5% percentiles), while the gray lines are for 11 weight combinations (2.5%, 10%, 20%, . . . , 80%, 90%, and 97.5% percentiles). The three lines for a given search are the 95% (dashed), median (solid), and 5% (dashed) summary lines from the top to the bottom, respectively. The step-function growth path summary shows the time to populate the front on a modern quad core processor desktop machine with the inherent variation from the random starts. Times for finding the target designs on the PF are summarized in Table 2. The median time to complete the search is around 2.7 min and the 90% range of time is between 0.74 and 9.52 min. This wide range of times indicates large variability of search time among the use of different random starting designs. Using 11 quantiles for determining the weight combinations is slightly faster for the median and 95% summary, but slightly slower than the 5 quantiles when the worst case scenario is considered. Generally the search guarantees finding all the target designs within 10 min for this uniform weight distribution scenario.

Fig. 5(b) shows the growth plot for a Beta (50, 50) distribution. Only two target designs (4 & 5) are sought for this distribution as the weight combination quantiles are centered around equal weights for the two criteria. The search finds the target designs much faster than for the uniform weighting preference with a median time of half a minute. The worst case (5% quantile) times are 0.87 and 1.21 min for 5 and 11 quantiles, respectively.

Fig. 5(c) and (d) show the Beta (15, 7) and Beta (71, 31) skewed weight distributions which place more emphasis on *D*-efficiency. The median time for finding the complete set of target designs is between 0.3 and 0.35 min for both distributions for either 5 or 11 quantiles. In this example, the efficiency of the search for both numbers of quantiles is similar, showing that there is considerable robustness to this choice. The richer search with more weight combination quantiles requires more time for each random start, but produces comparable results with fewer starts. The focused search with an appropriately specified weight distribution provides substantial improvements (about 2.8–9.3 times faster based on the median time, and 1.9–10.9 times faster for the 95th percentile for this example) with less variability compared to the complete front search.

Author's personal copy

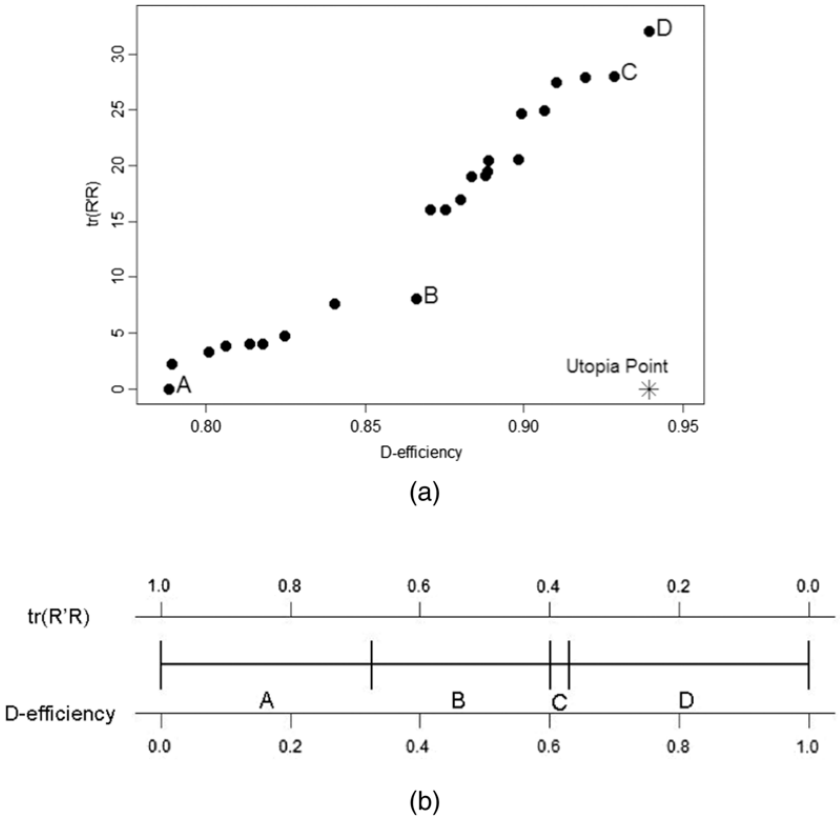


Fig. 6. (a) The complete PF based on *D*-efficiency and tr($\mathbf{R}'\mathbf{R}$); (b) best designs for different weightings of the two criteria for the additive DF.

Table 3
Criteria values for the 4 optimal designs for different weightings of *D*-efficiency and tr($\mathbf{R}'\mathbf{R}$).

Design	(*D*-eff., tr($\mathbf{R}'\mathbf{R}$))	Design	(*D*-eff., tr($\mathbf{R}'\mathbf{R}$))
A	(0.788, 0)	*C*	(0.928, 28)
B	(0.866, 8)	*D*	(0.939, 32)

Next we consider the same example with an alternative criterion, where an experimenter seeks a 14-run screening design for the same model and potential missing model terms as before, but now considering *D*-efficiency and tr($\mathbf{R}'\mathbf{R}$) criteria. In contrast to the tr($\mathbf{AA}'$) criterion, tr($\mathbf{R}'\mathbf{R}$) quantifies the potential bias for the estimate of experimental error variance from a mis-specified model, where $\mathbf{R} = \mathbf{X}_1\mathbf{A} - \mathbf{X}_2$ (see Myers et al. (2009)). Fig. 6(a) shows the richer PF with 23 designs obtained for the complete front and Fig. 6(b) shows the four optimal designs (*A*–*D* with criteria values shown in Table 3) for all weighting combinations when additive DF is used. Designs *A*, *B*, and *D* are all optimal for a wide range of weight combinations. Design *C* is optimal for only a very narrow range of weights when *D*-efficiency is weighted around 60%.

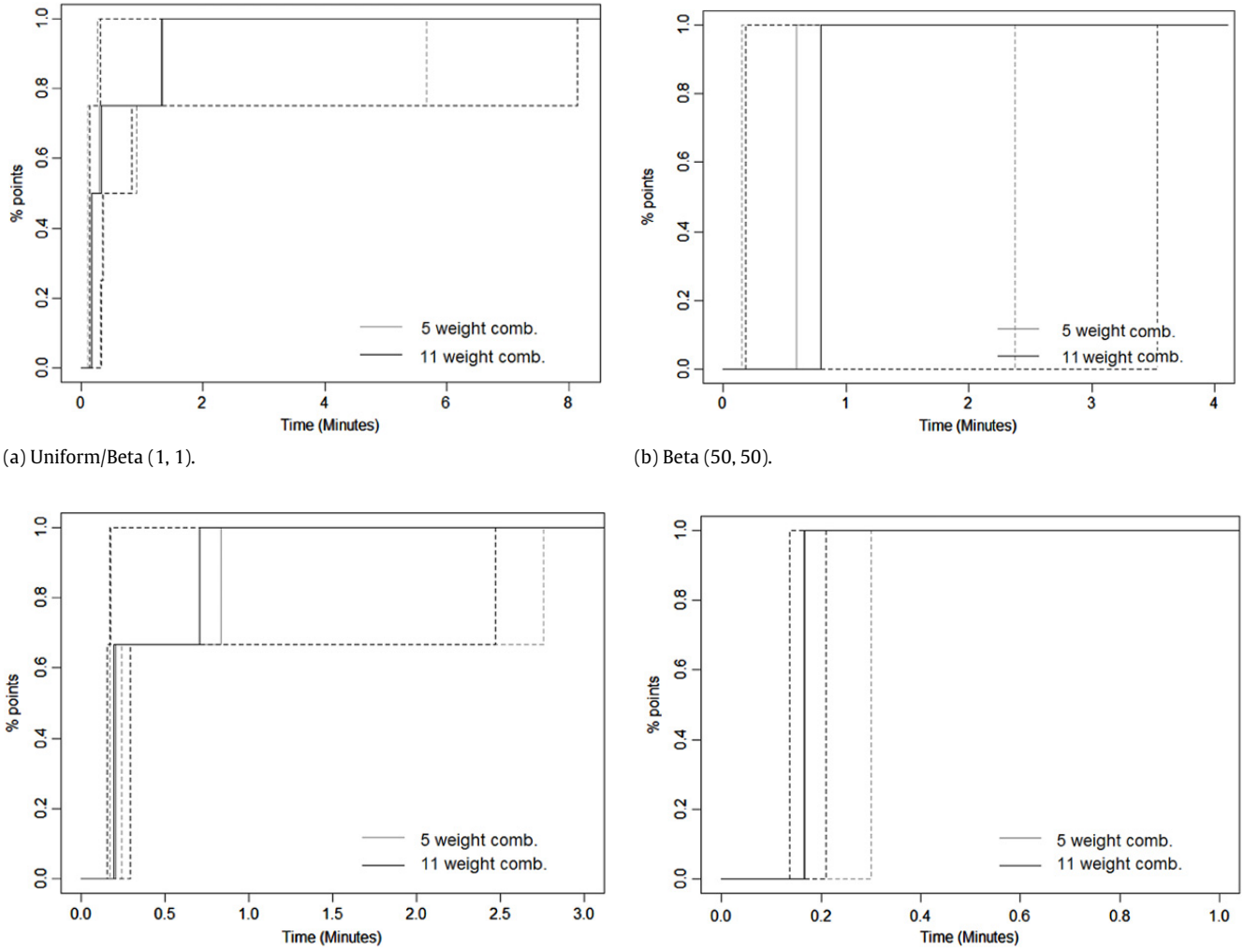
The same weight distributions from Fig. 3 are explored for assessing the performance of the focused PF search. Again the zero value on the desirability scale is selected to match the worst values from the complete PF. Table 4 shows the sets of target designs for the various weight distributions and the quantile summary of times for completing the search. Fig. 7 shows the front growth plot for different weight distributions, using the same parameters (N_b, N_r, N_s, and N_g) as the previous example. The median time for finding all four designs from the uniform distribution is around 1.3 min and 95% of searches are successful in no more than 8.2 min. When a more focused distribution is specified, the computing time for finding the target designs is substantially reduced with less variation in the process. Here, all the focused PF searches are more efficient (here, 1.2–7.9 times faster for the median time, and 1.8–38.8 times faster for the 95th percentile) with less variability.

Both examples demonstrate that the focused search can substantially reduce search times while still finding all the desired designs for decision-making based on focused preferences. If an alternate DF form is suggested by the SME, then this can easily be inserted into the search algorithm. The total times for the searches are relatively short for this problem involving a small design. However, as the number of runs and candidate locations increase, the total time for the PF search can grow quickly. Hence the focused PF search allows better scalability with even more advantageous time savings in these cases, and makes the PF approach more viable for a broader set of design of experiment situations. The PF growth plot effectively shows the timing for populating the front and quantifies the completeness, efficiency, and variation of certain search methods. The use of a resampling method for quantifying the process variation has computational advantages compared to running a large sample of independent searches for the complete front.

Author's personal copy

Table 4
Summary of computing times for finding the target designs on the PF based on *D*-efficiency and tr($\mathbf{R}'\mathbf{R}$) for different user-specified DF weight distributions for *D*-efficiency.

Weight distribution	Target designs	5 quantiles/weight comb.		11 quantiles/weight comb.	
		Median	95% range	Median	95% range
Uniform/Beta (1,1)	*A*, *B*, *C*, *D*	1.34	(0.26, 5.66)	1.32	(0.31, 8.14)
Beta (10,10)	*A*, *B*, *C*, *D*	1.15	(0.20, 3.12)	0.70	(0.17, 2.15)
Beta (50,50)	*B*	0.60	(0.15, 2.37)	0.80	(0.18, 3.53)
Beta (15,7)	*B*, *C*, *D*	0.84	(0.18, 2.76)	0.71	(0.17, 2.47)
Beta (71,31)	*C*, *D*	0.17	(0.14, 0.30)	0.17	(0.14, 0.21)



(a) Uniform/Beta (1, 1).

(b) Beta (50, 50).

(c) Beta (15, 7).

(d) Beta (71, 31).

Fig. 7. PF growth plot for different weight distributions for *D*-efficiency and tr($\mathbf{R}'\mathbf{R}$). The black and light gray lines represent searches using 5 and 11 quantiles for determining the fixed weight combinations. For each search, the three lines correspond to the 95% (dashed), 50% (solid) and 5% (dashed) empirical percentiles.

5. Decision-making based on a focused Pareto front

After obtaining the focused PF based on the user-specified distribution of weights of interest, the second stage is to incorporate the experimenter's priorities into decision-making by comparing the trade-offs and relative performance of the candidate solutions. If a non-uniform weighting preference has been specified for prioritizing the criteria, then this should also be utilized in this subjective phase.

We now return to the first example, where the experimenter has specified a Beta (15, 7) distribution for the weighting of *D*-efficiency, since it is considered more important than tr($\mathbf{AA}'$) in selecting the final solution. Note this weight distribution choice is dependent on a user's understanding of the problem and preference of the relative importance of the criteria. In eliciting this distribution, it may be helpful to have the user identify a point estimate for the primary weight combination

Author's personal copy

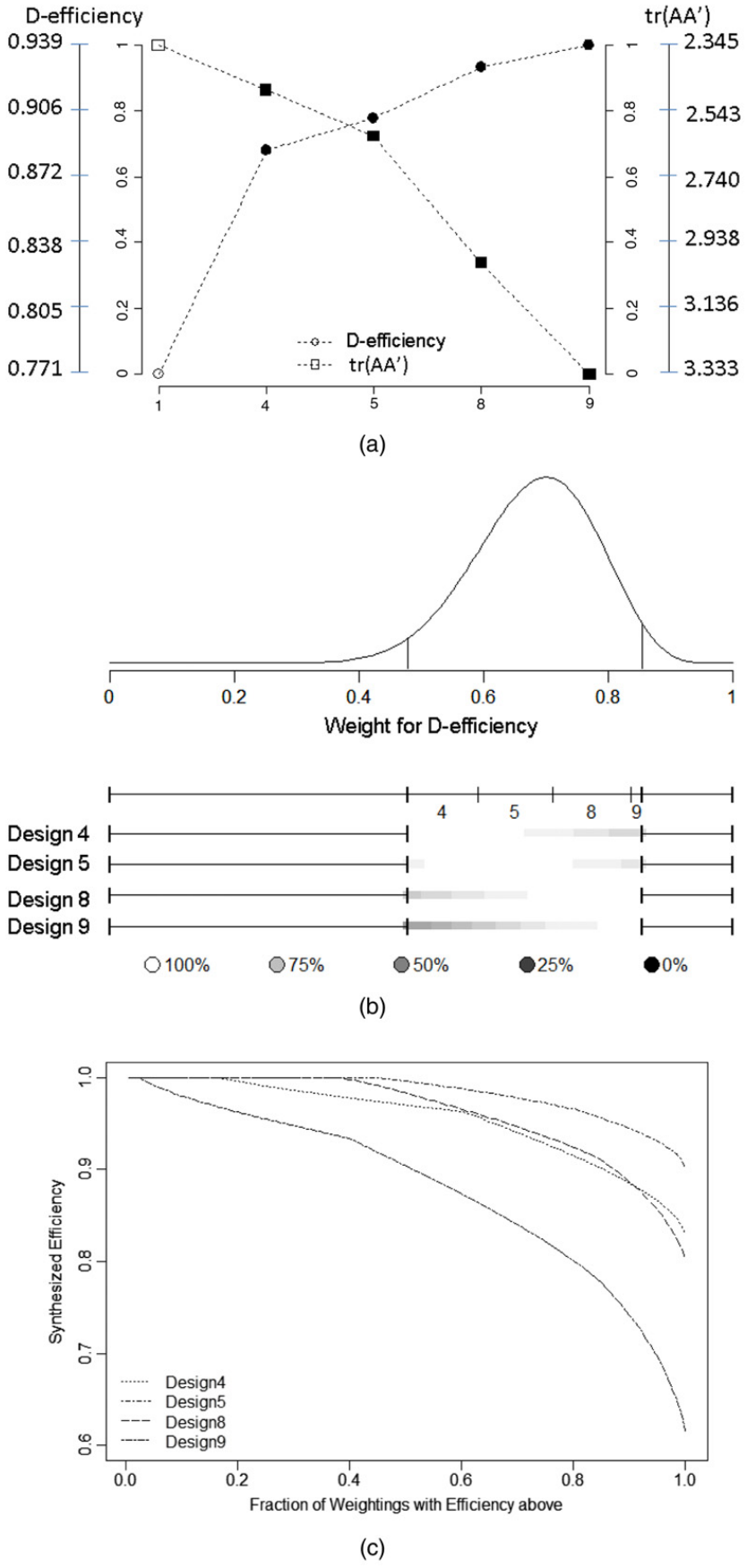


Fig. 8. Graphical summaries for a focused front based on *D*-efficiency and tr(**AA**′) for a user-specified Beta (15, 7) weight distribution: (a) trade-off plot; (b) best designs for different weightings of criteria and their synthesized efficiency plots; (c) FWS plot for designs selected in (b).

Author's personal copy

of interest, and then specify how much spread around that weight is of interest. If the user has limited knowledge on this, then a complete Pareto front search should be conducted without imposing any weighting preferences. Fig. 8 shows the graphical summaries adapted for the focused weighting preference. The focused PF search for Beta (15, 7) found 8 of the 9 designs from the complete front in Fig. 1. Design 2 was not found by the search, but this does not affect the final solution since it is not optimal for any weighting. With the additive DF being used, four (4, 5, 8, and 9) designs are in the focused range. The trade-off plot for the four target designs as well as the individual criterion optimal designs used to define the scaling of the plot are shown in Fig. 8(a). With the more focused weighting interest, the four designs are shown in black symbols. The tr($\mathbf{AA}'$)-optimal design 1 used to define the scaling is shown with an open symbol but is not of interest for the priorities of this study since it is optimal when tr($\mathbf{AA}'$) is weighted for at least 88%.

Fig. 8(b) shows the regions of different weightings corresponding to the four target designs within the focused range of interest and their synthesized efficiency plots. The top portion shows the specified weight distribution with the range of interest (the middle 95% area) bounded by the two vertical lines below the density curve. The best designs for different weightings within the range of interest are shown on the horizontal line segments underneath the weight distribution. This plot can help the experimenter identify the optimal design for any particular weighting of interest. Below are the synthesized efficiency plots for the four target designs. The plot shows the synthesized efficiency relative to the best possible design performance for every possible weightings in the range of interest. Note within the focused region, the weight distribution for optimal designs and the synthesized efficiency plots match results in Fig. 1(b) and (c).

Similar to Zahran et al. (2003) who suggest non-uniform weighting for the Fraction of Design Space plot for unequal emphasis of the design space, Fig. 8(c) shows the adapted FWS plot over the focused region for the four designs. Note that the FWS summary for each design is integrated using the weight distribution for the range of interest, and hence is dependent on the distribution shape and extreme percentiles selected by the user. For this example, the experimenter selects the 2.5% and 97.5% percentiles of the Beta (15, 7) to define the range. To calculate the FWS curve, a fine grid of evenly spread quantiles in the range is selected to determine weight combinations corresponding to the centers of a set of approximate rectangles under the distribution curve with approximately equal areas to equally partition the probability region. Summarizing with this set of weight combinations represents a discrete approximation to the user-specified continuous weight distribution within the region of interest. For the Beta (15, 7) distribution, design 5 has uniformly best efficiency across the focused weighting space, since design 5 is optimal for many weights with the highest density of the distribution. Design 8 has the second highest synthesized efficiency for the top 60% of the values, before its FWS line crosses the design 4 curve for a portion of the lower efficiency values. Design 9 has the uniformly worst synthesized efficiency over the entire region of interest, since it is best for only weights near the edge of the chosen distribution. Comparing Fig. 8(c) with Fig. 2 (where no design is uniformly best and design 4 is best for a majority of the weighting space), the relative performance of the four designs has changed substantially and design 5 becomes the global winner for the Beta (15, 7) preferred weighting distribution.

To illustrate the impact of the distribution within the range, consider the same range of interest, but with a uniform distribution in Fig. 9. In this case, weights around the peak area of the Beta (15, 7) distribution are considered equally important to the boundary region. Comparing the results with Fig. 8(c), we see that design 5 is still the universal winner, but the proportion of efficiencies above various thresholds becomes slightly smaller since the high efficiency area is weighted less for this distribution. Design 8 is better than design 4 for the upper 40% efficiency of interest. Otherwise, design 4 is the second best across a majority of the weight space considered. Design 9 still has the worst synthesized efficiency across the focused region. As is desirable, different user emphases within the region can result in different relative performance of designs as summarized by FWS, and hence could potentially lead to different final decisions. This reinforces the benefits of this approach as the specific priorities of the experimenter can be quantitatively incorporated through the user-specified distribution to reach a tailored solution to the experiment of interest. It is also straightforward to examine the sensitivity of results to different distributions within a given range.

6. Discussion and conclusions

The PF approach can guide practitioners to improved decisions when balancing multiple objectives. The objective PF search process identifies contending choices considering all criteria simultaneously. Then an informed decision can be made based on considering the trade-offs between alternative choices and their performance for different emphases of the criteria that match the goals of the study. The method allows great flexibility without much extra computational work for evaluating the impacts on the chosen solution of subjective aspects of the decision-making including the prioritization/weighting of the criteria, scaling used for converting the criteria values onto a 0–1 scale, and the metric for integrating multiple criteria into an overall summary index.

Finding all contending solutions in the objective search stage is important for understanding what alternatives exist, and to calibrate the possible range of criteria values. However, the complete PF search can be time consuming for larger design sizes, many candidate design locations and more criteria. To facilitate broader adoption of the PF approach for selecting tailored designed experiments, the development of efficient search algorithms with good scalability is important. This paper focuses on improving the computational efficiency of the search when *a priori* there is a focused preference on the prioritization of the criteria. This is a common scenario for many design selection problems where a region of possible criterion weightings can be specified by a subject matter expert or a decision-maker with a non-uniform distribution to

Author's personal copy

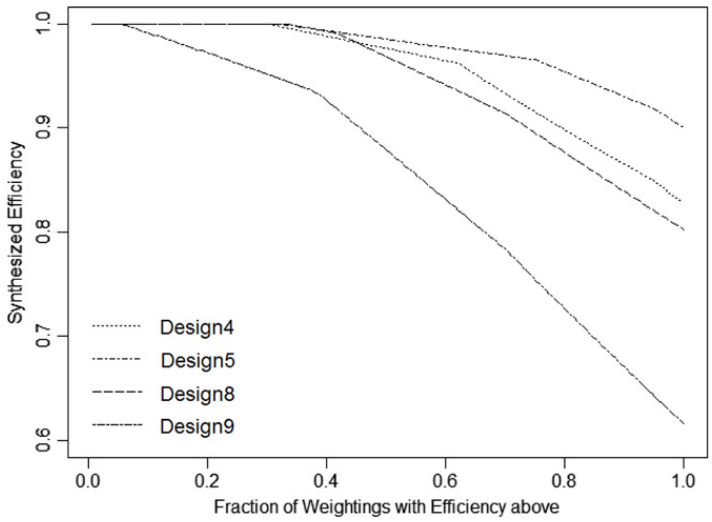


Fig. 9. The FWS plot obtained by using a uniform distribution for the 95% middle interval of weights for a focused PF based on *D*-efficiency and tr(**AA**′) with a user-specified Beta (15, 7) distribution.

summarize the unequal preference on different weightings. For a user-specified weight distribution, we develop the focused PF search algorithm using fixed weightings (Lu and Anderson-Cook, 2012) to improve the search efficiency. Multiple fixed weight combinations within the range of interest determined by a set of quantiles of the specified distribution are used to guide the search heading in diverse directions simultaneously for better coverage, efficiency and consistency of performance.

At each updating step, the DF indices are calculated for each set of the fixed weights for both the current and new designs generated with the modified PAPE algorithm. To use the focused PF approach, some decisions are required. First, the criteria values need to be converted to a 0–1 scale to calculate the overall DF index. Hence best and worst values must be specified before performing the search. The best value can be found easily by conducting a single criterion search. The worst value can be specified based on either using a natural worst value or from preliminary searches based on subsets of the criteria. When using a set of well spread multiple weight combinations, the search result is somewhat robust to different scaling choices.

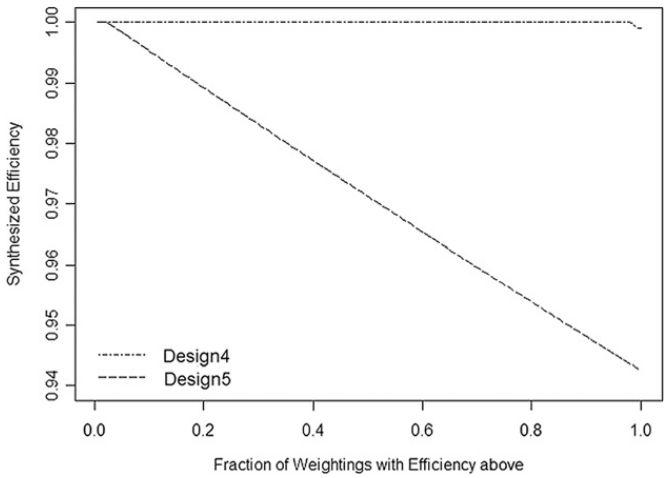
Second, other user-specified choices are the set of quantiles used for determining the extreme percentiles for the range to be explored and set of weight combinations for directing the search. A finer set of weights provides good coverage of the front, but does increase the computational time as the number of parallel searches increases. In the screening design example in Section 4, using 5 and 11 quantiles are explored. The use of 11 quantiles has slightly better performance for most of the scenarios explored, but with little practical difference observed.

Third, a DF metric should be selected for integrating multiple criteria into an overall summary. The additive and multiplicative DFs are two common options with different penalties on worst performance for at least one of the criteria. Other forms have also been suggested in the literature. The experimenter should decide which metric is most appropriate to summarize study goals. The example in Section 4 used the additive DF, but could easily be adapted to other metrics. Note that the metrics used during the searching process and for the graphical summaries should be consistent and match the experimenter priorities.

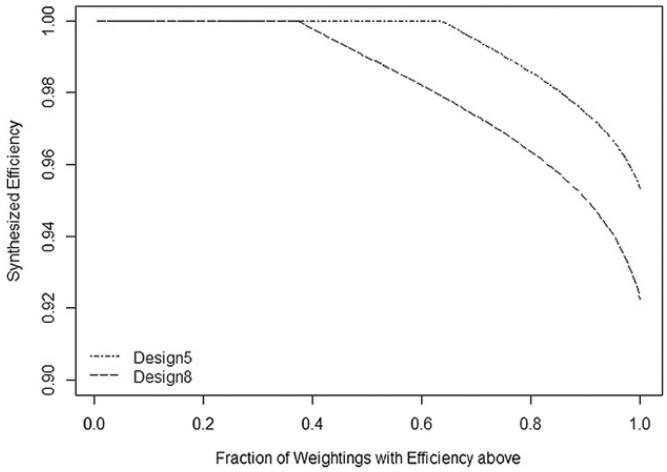
Two alternative updating mechanisms with comparable performance, using fixed or stratified weightings, are proposed in Lu and Anderson-Cook (2012) for improving the search efficiency. This paper demonstrates substantial computational gains from using the more focused PF search method with fixed weightings. Adapting the focused PF search to use stratified random weightings is straightforward. This ensures parallel searches are conducted in diverse directions with small variation allowed at each updating step for a given stratum to incorporate the weight specification uncertainty into the search process. The performance of the search methods can be accessed and compared using the PF growth plot in Section 4, where the performance is evaluated based on the proportion of optimal designs found within the range of interest with the process variation accessed using a resampling method.

After obtaining the PF with the focused search algorithm, adapted graphical summaries can identify good solutions that reflect the user preferences. The adapted FWS plot provides an overall quantitative summary of individual design performance based on the priorities of the experiment and allows easy comparison between design choices. The FWS plot reflects not only the range of weights of interest, but also their relative emphasis based on the user-specified weight distribution. Fig. 10 shows the FWS plot for using two other weight distributions, (a) Beta (50, 50) and (b) Beta (71, 31), which differ from Figs. 8(c) and 9. For the centered Beta (50, 50) distribution, design 4 is the universal winner across all possible weightings. However, for a skewed Beta (71, 31) distribution with more weights on *D*-efficiency, design 5 is the best choice across the entire weighting space. Therefore, carefully specifying the weighting distribution to capture the experimenter's prioritization of multiple criteria is important for finding the right solution. With minimal additional computational effort,

Author's personal copy



(a) Beta (50, 50).



(b) Beta (71, 31).

Fig. 10. FWS plot for target designs from the PF based on *D*-efficiency and tr(**AA**′) for different user-specified weight distributions: (a) Beta (50, 50) and (b) Beta (71, 31).

a sensitivity analysis can easily examine the impact from different weight distributions within a fixed range of weighting interest.

References

Coello Coello, C., Van Veldhuizen, D., Lamont, G., 2007. Evolutionary Algorithms for Solving Multi-Objective Problems, second ed. Kluwer Academic Publishers, New York.
Deb, K., 2001. Multi-Objective Optimization Using Evolutionary Algorithms. John Wiley & Sons, Chichester, UK.
Derringer, G., Suich, R., 1980. Simultaneous optimization of several response variable. Journal of Quality Technology 12, 214–219.
Gronwald, W., Hohm, T., Hoffmann, D., 2008. Evolutionary Pareto-optimization of stably folding peptides. BMC-Bioinformatics 9, 109.
Harrington, J., 1965. The desirability function. Industrial Quality Control 21 (10), 494–498.
Kasprzak, E.M., Lewis, K.E., 2001. Pareto analysis in multiobjective optimization using the collinearity theorem and scaling method. Structural Multidisciplinary Optimization 22, 208–218.
Lu, L., Anderson-Cook, C.M., 2012. Rethinking the optimal response surface design for a first-order model with two-factor interactions, when protecting against curvature. Quality Engineering 24, 404–422.
Lu, L., Anderson-Cook, C.M., 2012. Adapting the hypervolume quality indicator to quantify trade-offs and search efficiency for multiple criteria decision-making using Pareto fronts. Quality and Reliability Engineering International http://dx.doi.org/10.1002/qre.1476.
Lu, L., Anderson-Cook, C.M., Robinson, T.J., 2011. Optimization of designed experiments based on multiple criteria utilizing a Pareto frontier. Technometrics 53, 353–365.
Lu, L., Anderson-Cook, C.M., Robinson, T.J., 2012. A case study to demonstrate Pareto frontiers for selecting a best response surface design with simultaneously optimizing multiple criteria. Applied Stochastic Models in Business and Industry 28, 206–221.
Lu, L., Chapman, J., Anderson-Cook, C.M., 2013. Selecting a best allocation of new data for improving the estimation precision of system and subsystem reliability using Pareto fronts. Techometrics (in press).
Myers, R.H., Montgomery, D.C., Anderson-Cook, C.M., 2009. Response Surface Methodology: Process and Product Optimization Using Designed Experiments, third ed. Wiley, New York.
Trautmann, H., Mehnen, J., 2009. Preference-based Pareto optimization in certain and noisy environments. Engineering Optimization 41, 23–38.
Wagner, T., Trautmann, H., 2010. Integration of preferences in hypervolume-based multiobjective evolutionary algorithms by means of desirability functions. IEEE Transactions on Evolutionary Computation 14 (5), 688–701.
Zahran, A., Anderson-Cook, C.M., Myers, R.H., 2003. Fraction of design space to assess the prediction capability of response surface designs. Journal of Quality Technology 35, 377–386.