# Multi-Criteria-Optimisation and Desirability Indices

Dipl.-Stat. Detlef Steuer
Universität Dortmund
LS Computergestützte Statistik
steuer@amadeus.statistik.uni-dortmund.de

March 1999

**Abstract**

The basic ideas of Desirability functions and indices are introduced and compared to other methods of multivariate optimisation. It is shown, that gradient based techniques are not in general appropriate to perform the numerical optimisation for Desirability indices. The problems are shown for direct modelling of Desirability indices. An example is given to illustrate the sensitivity of estimated optimum factor settings to modelling errors for individual targets.
Keywords: MCO, MCD, MCDA, Desirability Function, Desirability Index, dimension reduction, numerical optimisation.

## 1 Introduction

When dealing with industrial production processes we often have to assess quality of products. It turns out, that often more than one measured variable or property has to be taken into account to describe "quality". Normally quality not only should be measured, but also tried to improve.
The optimisation of quality will normally prove more difficult in the case of two or more competing properties. Often it is not possible to improve one of them without deteriorating one or more of the others. In the context of optimisation properties are also called "targets".
A very common example can be found in the context of drug-design: A new pill has to fulfil two, often contradictory, requirements. For example firstly it should be very well solvable in water, secondly it should not fall apart without

water. While experimenting, the designer realizes that improving solvability deteriorates durability. Solvability and durability here are two competing properties in drug-design. The problem with competing properties shows up, if there exist two designs, one with better solvability, the other with better durability. It becomes impossible to rank these alternatives without knowledge about the relative importance of the variables. Weights of some kind have to be introduced.
The described situation of multiple, partly contradictory objectives in optimisation is the starting point of a branch of Operations Research (OR), called Multi-Criteria-Optimisation (MCO), Multi-Criteria-Decision-Making (MCD) or Multi-Criteria-Decision-Analysis (MCDA). It is clear, that problems of this kind are not restricted to industrial production processes. They appear wherever the "value" of a "decision" is measured in more than one dimension. The word "product" will be used in this general interpretation throughout the paper.
For this paper a product's multiple properties are assumed to be competing, i.e. they can not be optimised at the same time. Otherwise all problems would reduce to the one-dimensional case of optimisation, which is not of interest here.
The aim of this work is to discuss some of the OR methods to handle the situation of multiple objectives from a statistical point of view, with stress on the so called desirability indices.
In Section 2 the formal context of multiple objectives from a statistical point of view will be introduced, where the properties can not be set directly, as it is often assumed in OR (within some restricted area). Instead properties have to be understood as functions of some underlying controlled variables (factors) which must be approximated via statistical modelling. Section 3 presents mathematical and practical concepts for MCO. The concept of "desirability" is explained in Section 4. After that two examples for "desirability functions" used in practice are given in Sections 5 and 6. In Section 7 we show, that gradient based optimisation techniques are unsuitable for the presented optimisation problem without further considerations. We discuss some aspects of modelling "desirability indices" in Section 8 and give a short conclusion in Section 9.

## 2 Formalism

In this section the formal setting of this paper is defined.
The focus here lies on a vector $P \in \mathbb{R}^Z$, which is defined by its coordinates $Y_i(P) = p_i, i = 1, \ldots, Z$. This vector $P$ stands for a product with $Z$

properties, which describe its quality.
The properties $Y_i, i = 1, \dots, Z$, are functions of a finite number of factors $X_1, X_2, \dots, X_F$ and a stochastic error term $\epsilon$. Therefore we get:

$$P = \begin{pmatrix} Y_1(P) \\ Y_2(P) \\ \vdots \\ Y_Z(P) \end{pmatrix} \text{ with } Y_i(P) = f_i(X_1, \dots, X_F, \epsilon), i = 1, \dots, Z.$$

This formula deserves some remarks:
First the same set of factors is used for all properties in this formula. This is not a restriction, if we allow them to be purely formal parameters of the $f_i$. Second we note the error is not restricted to be additive!
Under these assumptions the objective in MCO in this statistical context may be formulated as finding the setting $X_{opt} = (X_{1,opt}, \dots, X_{F,opt})$, for which the expectation $E(Y_1, Y_2, \dots, Y_Z)$ is "best"!
Implicitly this calls for a ranking in $\mathbb{R}^F, F > 1$. As is well known, no such ranking exists. Therefore it can only be tried to get "close to a ranking" in some sense. In the following the basic concepts of multivariate ranking will be shown.

# 3 Mathematical concepts for MCO

In this section mathematical concepts and problems for ranking multivariate observations will be presented. These considerations lead to requirements for practical MCO-procedures.

## 3.1 Domination

A first, very optimistic, approach to MCO is hoping for one "really best" object, which means an object that is better than all alternatives in every of the multiple targets. The idea of a ranking in $\mathbb{R}^Z$ is abandoned here. The problem of *ranking* $\mathbb{R}^Z$ is *reduced to finding the "maximum"* in a given set of alternatives using a coordinate wise "better" relation, symbolised here by ">".
This concept is called domination and a formal definition is given below:

**Definition 1 (Domination)** *Given two objects $P_{1,2} \in \mathbb{R}^Z$ described by the same properties $Y_i(P)$, $i = 1, \dots, Z$, it is said, that $P_1$* **dominates** *$P_2$ ($P_1 >> P_2$), if $Y_{i1} \geq Y_{i2} \quad \forall\, i = 1, 2, \dots, Z$ and $Y_{j1} > Y_{j2}$ for one $1 \leq j \leq Z$.*

At the same time the domination of factor settings has been induced as follows:

**Definition 2 (Domination in factor space)** *Given two settings $X_1$ resulting in product $P_1 := E(P(X_1))$ and $X_2$ resulting in product $P_2 := E(P(X_2))$ and the products described by the same properties as above, it is said, that $X_1$* **dominates** *$X_2$ ($X_1 >> X_2$), if $P_1 >> P_2$.*

Formally this is a nice and clean mathematical concept. For practical use it has an important drawback: There may not exist a dominating object. And worse, as $Z$ increases, it becomes more and more unlikely for an object to dominate any other.

## 3.2 Pareto-Optimality

Having seen the probable non-existence of solutions using domination, the less strict concept of "pareto-optimality" can be tried. It is defined as follows:

**Definition 3 (Pareto-Optimality)** *Given a set $M \subset \mathbb{R}^Z$ of objects, an object $P \in M$ is called* pareto-optimum*, if there is no object $Q \in M$ with $Q >> P$.*

Or less formal: If an object can not be improved in the coordinate wise sense of domination, it is optimal in some way. Inversely, and perhaps more important: If an object is not pareto-optimum, it should not be considered as a solution of a MCO problem!
Pareto-Optimality in factor space can be defined analogous to domination in factor space (see definition 2):
A factor setting $X$ is pareto-optimum in factor space, if the corresponding product $P := E(X)$ is pareto-optimum in product-space.
This leads to a first strict requirement for a MCO procedure: Any proposed solution of a MCO problem must be pareto-optimal!
While the problem with domination is the missing guarantee for finding a solution, the problem with pareto-optimality is -in general- the possibility of many solutions for the MCO.
On the other hand the experimenter is looking for a unique, best solution, so it has to be decided among the proposed solutions. To accomplish a decision weights for the different targets must be introduced, representing their *relative* importance. The final decision will be highly problem specific. Thus a second requirement for MCO procedures is found: it has to be possible to take different importances of targets into account!

In a mathematical sense we are looking for a functional $C : \mathbb{R}^Z \to \mathbb{R}$, which allows the ranking of all objects in $\mathbb{R}$ in an appropriate, problem specific way.

### 3.3 Approaches to MCO in OR

A lot of heuristic methods for solving MCO problems have been developed in OR. The statistical properties of these algorithms often are not well known. This is not too surprising, as these algorithms were not introduced with focus on their statistical aspects. Instead they are constructed to generate solutions in practice.
A short list of approaches symbolising different classes of MCO-solvers, could contain the following entries:

- Pareto optimality, standing for the mathematical, not problem specific approach,
- overlay plots, a graphical method,
- Prometee, representing the so called outranking procedures,
- desirability functions and -indices, which will be presented in detail in this paper and
- (monetary) loss functions, a member of the utility function class.

Each of these entries has its specific advantages and disadvantages. Pareto optimality has been considered in the section above. Overlay plots try to simultaneously look on contour plots of all targets. This method fails with more than three or four targets. Prometee uses pairwise comparison of all possible experimental results for constructing a (partial pre-) order among them. For each target $Y$ a so called preference function $P_Y$ is defined, that assigns a numerical value to the difference of two experiments $a$ and $b$ in this target: $P_Y(a, b) = P_Y(a-b)$. An expert defined weighted sum of these target preferences serves as "preference index" $\Pi(a, b)$ for comparison of two experiments. The sum over all preference indices $\sum_b \Pi(a, b)$ is called the "positive outranking flow" for experiment $a$ then. The sum $\sum_b \Pi(b, a)$ is called the "negative outranking flow" of experiment $a$. Based on these "flows" a ranking of all experiments can be produced. Prometee as a MCO tool is very flexible to use, but hard to interpret. Little is known about its robustness.
For monetary loss functions there is a problem in their concreteness: It is in general not possible to specify them as exact as it would be required to really interpret them as money value.

A more detailed overview about these techniques can be found in [4]. Desirability functions will be explored in the next sections.

# 4 The Concept of Desirability

In this section we will describe the basic concept of desirability for MCO.
In the formal description of the MCO problem we find $Z$ different targets to be optimised. Translated to real world problems it means that up to $Z$ different scales of measurement, i.e. times, lengths, weights and so on, have to be forced to be commensurable. Some of the existing heuristics try to ignore this problem. They assign a weight to each target and compare the weighted units.
Desirability is different here. In a first step every objective $Y_i, i = 1, \ldots, Z$ gets translated by an individual, so called "desirability-function" into a unitless desirability-scale. The step of defining the desirability for every possible outcome of $Y_i$ obviously becomes crucial to the process of multivariate optimisation. Naturally the statistician is unable to determine good or bad weights! A close collaboration of statistician and expert is very important when fixing the desirability function for a target.
Which are useful functions for using as desirability-functions $d$? First we can force them to take values in $[0, 1]$. The limitation to $[0, 1]$ has a technical justification, which will become obviously later on.
The most general definition of a desirability function now is the following:

**Definition 4 (Desirability function)** *Any function $d$ with*

$$d : \textit{Domain of } Y \to [0, 1], Y \hookrightarrow d(Y).$$

*is called* **desirability function (DF)**.

Not all of these functions may usefully serve as DFs. To fulfil the requirements formulated in Section 3 it has to be required, that

1. $d$ is flexible enough to allow problem specific formulations. It should be possible to give parameters $LSL$ (lower specification limit), $USL$ (upper specification limit), $T$ (target value) and, if needed, possibly different weights $\beta_l$ and $\beta_r$ for deviations to the left respectively right of $T$.

2. For target value problems $d$ should increase monotonically in $(-\infty, T)$ and decrease monotonically in $(T, \infty)$. This guarantees the pareto-optimality of any desirability-optimum solution of the optimisation problem!

Perhaps the last property requires a little proof:

**Lemma 1 (Proof: Desirability optimum points in factor space are pareto-optimum.)**
A desirability optimum $X_{opt}$ is a local optimum of the desirability-function over factor space. Assume $X_{opt}$ is not a pareto-optimum point. Then there exists $Y_{opt}$ in factor space, which gives a superior result and therefore higher desirability. Contradiction! $\square$

Concrete examples of DFs used in practice are given in Sections 5 and 6.
After scaling all targets individually to a desirability scale, they have to be combined to a single number, the overall desirability or desirability index of the product $P$. The following gives the general formal definition for a desirability index:

**Definition 5 (Desirability Index)** *A desirability index* **DI** *is a function $D$ with*

$$D : [0,1]^Z \to [0,1].$$

The obvious idea for a concrete formulation of D is to use some kind of mean value. Often the geometric mean is used for $D$. It has the feature to assign a DI of 0, if any of the individual DFs is 0. This is nicely interpretable: If one of the product's properties is completely unacceptable, the product as a whole is unacceptable. The use of the geometric mean also gives justification for choosing the interval $[0,1]$. Individual desirabilities greater than 1 would allow to compensate shortcomings in some of the other properties. Nevertheless other function are used also. Of these the maximin DI is most important. It defines

$$D(P) := \max_{\mathbf{X}} \min_{i=1,\ldots,Z} d_i(\mathbf{X}).$$

This formulation can be nicely interpreted also: A product is only as good as it's worst property at the optimum factor setting.
In the statistical context the desirability index $D$ may now be represented as a function of the factors $X_{i1}, \ldots, X_{iF}$ and the unknown errors $\epsilon_i,\ i = 1, \ldots, Z$.

$$\begin{aligned} D(P) &:= D(d_i(Y_i)_{i=1,\ldots,Z}) \\ &= D(d_i(f_i(X_{i1}, X_{i2}, \ldots, X_{iF}, \epsilon_i))_{i=1,\ldots,Z}) \end{aligned}$$

Each factor setting is evaluated by a number $D \in I\!R$. The canonical ranking in $I\!R$ induces a ranking in $I\!R^Z$ and the "best" factor setting can easily be

identified. As a result the multi objective optimisation problem has been turned into a response surface problem.
In literature about applications of DFs the most common forms of DFs are those of Harrington and Derringer/Suich. Both will be presented here with their formulation of desirability functions and indices for target value problems.

# 5 Desirability Function I

The concept and the name "desirability" were introduced by Edwin C. Harrington Jr. ([3]) in 1965.
For DFs Harrington used the following exponentials to handle target value problems:

$$d^H(Y) := e^{-|Y'|^n}, \text{ where } Y' \text{ is an appropriate transformation of } Y. \quad (1)$$

Appropriate in the sense of Harrington is to choose $Y'$ in a way, so that $d^H(LSL) = d^H(USL) = 1/e$. As a possible transformation he gives

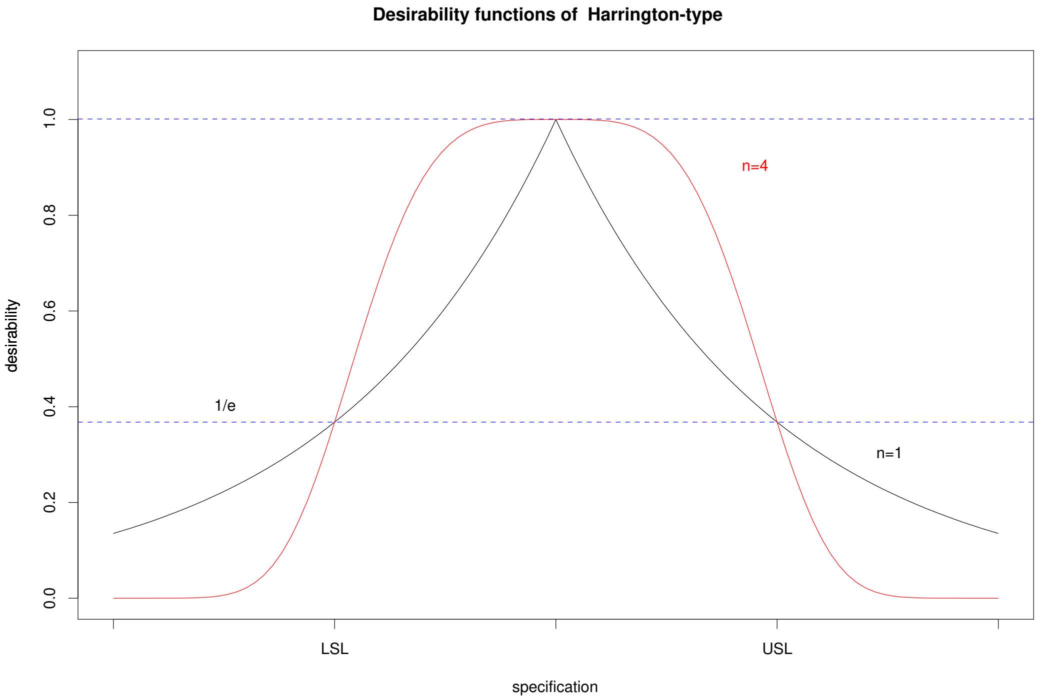
$$Y' = \frac{2Y - (USL + LSL)}{USL - LSL}.$$

Defined this way $d^H(Y)$ is symmetric around the centre between LSL and USL. The parameter $n$ serves as an "deviation importance" parameter. Large values of $n$ result in flat curves around the centre, thus punishing deviations from the centre less hard than low values of $n$. Figure 1 shows two typical DFs of this type, with parameters $n = 4$ and $n = 1$.
For combining all $d_i^H,\ i = 1, 2, \ldots, Z$ Harrington preferred the geometric mean. Written as a function in $Y_i,\ i = 1, 2, \ldots, Z$ and finally in $X_j, j = 1, 2, \ldots, F$ we get the DI of Harrington as:

$$\begin{aligned} D^H(P) &:= \sqrt[z]{\prod_1^Z d_i^H(Y_i)} \\ &= \sqrt[z]{\prod_1^Z d_i^H(f_i(X_1, X_2, \ldots, X_F, \epsilon_i))}. \end{aligned}$$

For optimisation purposes the Z-th root is unimportant, however for comparing the overall desirability with the per target desirabilities it is essential. Harrington developed his DF with interpretation of the result in mind. In his paper he even gave a scale for interpretation to use with his index. He proposes to interpret

Figure 1: Desirability functions of Harrington-type for two different values of n




- desirability 1 as "ultimate satisfaction" or "improvement beyond this point has no value";
- desirability $0.8-1$ as "excellent" or "well beyond anything available";
- desirability $0.63-0.8$ as "good" or "slight improvement over industrial quality";
- desirability $0.4-0.63$ as "acceptable, but poor";
- desirability $0.3-0.4$ as "borderline";
- desirability $0-0.3$ as "unacceptable to completely unacceptable".

There are some disadvantages to Harrington's approach. The DFs are not too flexible, as for example they are always symmetric. Furthermore Harrington has to use different exponential functions for maximisation (minimisation) problems, thus hardening the comparability of multiple desirabilities.

On the other hand his functions have a big plus for being given in closed form and being differentiable.

# 6 Desirability Function II

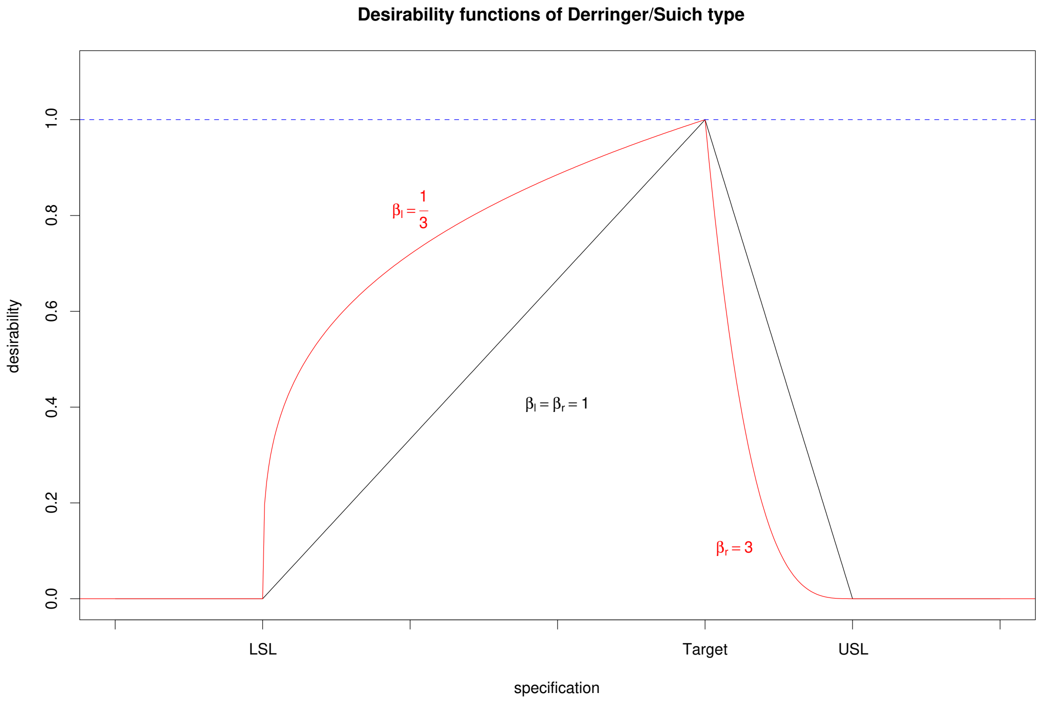
Derringer and Suich defined a new class of desirability functions in [2] to gain flexibility for modelling the importance of individual targets.
Their definition of a desirability function for target value problems is given below:

$$d^{DS}(Y) := \begin{cases} 0, & \text{for} \quad Y < LSL \\ (\frac{Y - LSL}{T - LSL})^{\beta_l}, & \text{for} \quad LSL \leq Y \leq T \\ (\frac{USL - Y}{USL - T})^{\beta_r}, & \text{for} \quad T < Y \leq USL \\ 0, & \text{for} \quad USL < Y \end{cases}$$

The parameters $\beta_l$ and $\beta_r$ are weights for deviations to the left respectively to the right from the target. Values near 0 mean unimportant deviations, while high values stand for very important targets. In figure 2 characteristic desirability functions have been plotted, to show the flexibility of this new DF class.
For constructing the DI Derringer/Suich also propose the geometric mean.
Mini- or maximisation problems can be handled consistently by using only one branch of $d^{DS}$. In this case $T$ defines a value for which all lower (higher) values are accepted as "perfect", thus giving desirability 1. Besides the possibility to model asymmetry herein lays the main advantage of Derringer/Suich type functions over those of Harrington type. Furthermore it is straightforward to generalise these functions to more than two segments.
However, they could not get the desired flexibility for no price: the closed form of the DFs had to be changed to a piecewise definition. On the plus side we find unacceptable target values having desirability 0, thus giving overall desirability 0 if a single target is unacceptable. The parameters $\beta_l$ and $\beta_r$ are important for the optimisation step, so they have to be chosen carefully. Optically there seems to be an easy interpretation of the different values of $\beta_{l,r}$. In practice a doubled value of $\beta$ is told to give a doubled importance to deviations from the target value. Mathematically it is obvious that this simplicity is only a "rule of thumb" and has to be questioned.
Due to the superior flexibility of the Derringer/Suich function class Harrington's approach could be discarded. But piecewise definition with non-differentiable points at interval borders lead to problems in analytical handling of the new DFs.

An implementation of Derringer/Suich desirabilities is found in STAVEX, a software package for design and analysis of experiments, developed at CIBA-GEIGY, Basel [5].

Figure 2: Desirability functions of Derringer/Suich type for two different values of $\beta_l$ and $\beta_r$, asymmetric case.

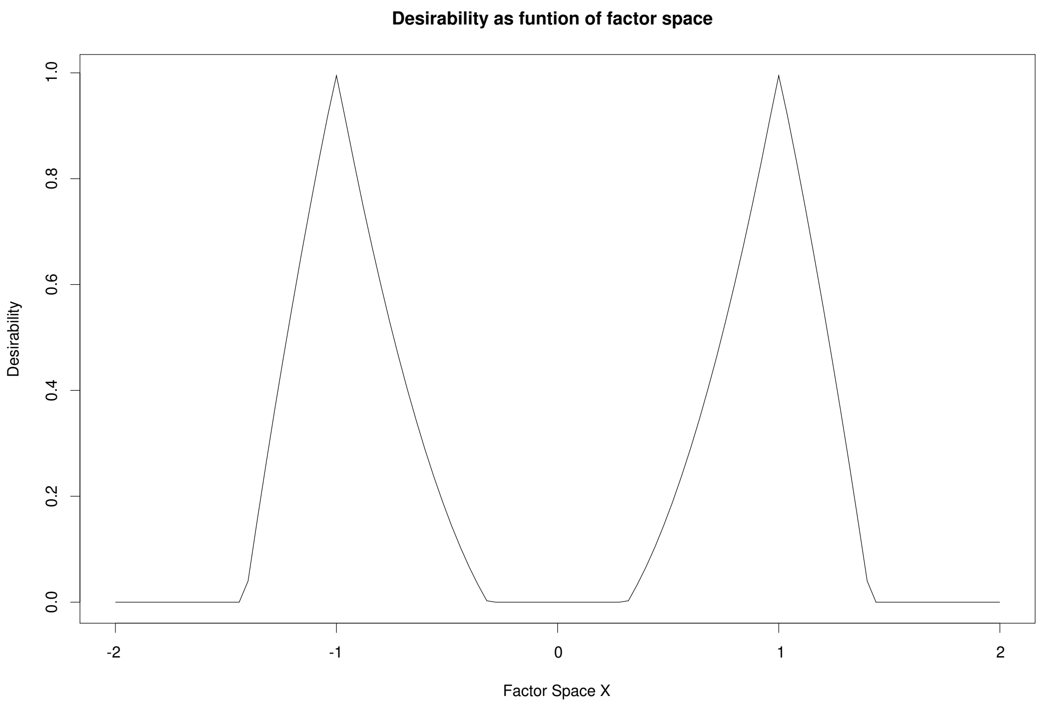



# 7 Practical optimisation issues

The main achievement using desirability indices has been the conversion of the original problem into a response surface problem. The founders of the DFs presented here wanted to use the well known optimisation techniques for response surfaces, mainly the gradient based ones. For this purpose the Derringer/Suich DFs had to be refined to be differentiable everywhere. That work was done by [1] by simply fitting fourth degree polynomials around each non differentiable point.
In principle both types of DFs allow the usage of classical techniques for optimising a real-valued function. In literature those techniques were applied without any further considerations. See for example [2] or [1]. This

behaviour by the scientists may have been supported by the apparently simple structure of the DFs. Each DF has a single optimum, which seems to assure unimodality of the respective DI. Unfortunately this turns out to be an oversimplification.

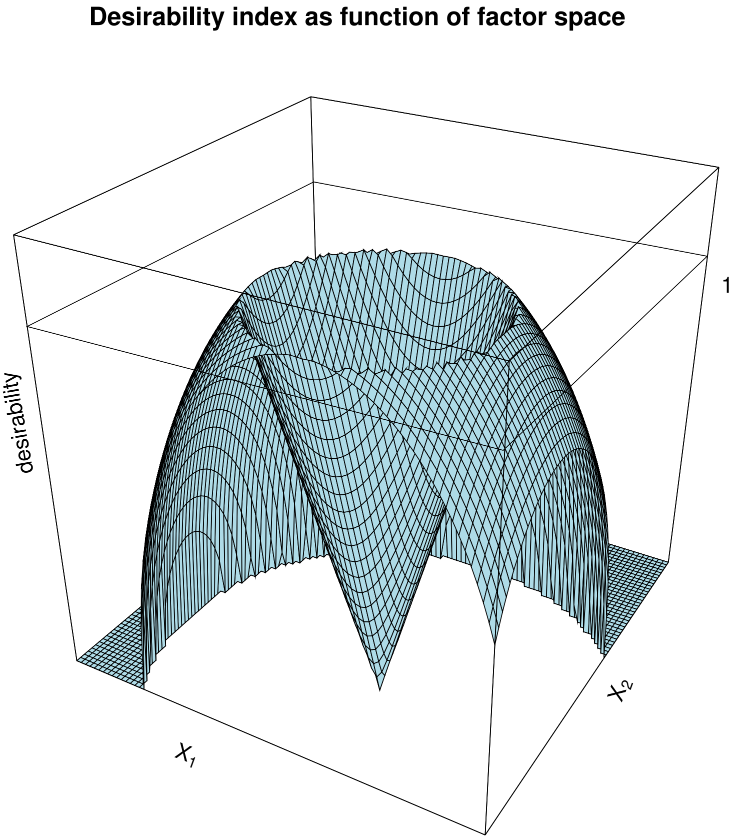
Figure 3: Desirability function of Derringer/Suich type for $Y = -X^2+1$; $T = 0.5$; $LSL = -1$; $USL = 1$; $\beta_l = 1$; $\beta_r = 1$ as function of factor X (1-dim)




If considered as mappings from factor space into desirability scale DIs are not unimodal! Figure 3 gives the most simple example for bimodality of a DI as a function of a single controllable variable. In this special case the DF equals the DI, because only one target exists. As Figure 3 shows factor space can even fall apart in disjunct regions of possible desirability. Furthermore more than one local maximum may exist. It can easily be seen, that not all local maxima of a DI must be of same height, if more than one target is considered.
As a consequence classical (gradient based) optimisation techniques may fail, if their starting point was not chosen very well. After looking at Figure 3 this is obvious, but it was not considered by many experts and was not mentioned in literature known to us. Before any gradient based technique can be applied

to optimise a DI, unimodality of the response surface has to be assured.

Figure 4: Desirability function of Derringer/Suich type as a function of factor space (2-dim). $Y = X_1^2 + X_2^2$, LSL=-1, T=0, USL=4, $\beta_l = \beta_r = \frac{1}{2}$.




An alternative approach to optimise DIs is used in STAVEX [5]. An exhaustive grid search is performed to find the optimum settings. This approach is independent of unimodality of the response, but computing time restricts this approach to low dimensions of factor space. The cause for using grid search in STAVEX has been the non differentiability of the Derringer/Suich DFs. A modern implementation should use a sophisticated search heuristic like simulated annealing for DI-optimisation in the case unimodality can not be assumed.
To give an impression of the beauty and the possible complexity of even a single DF the example in Figure 4 is given. In there a DF of a single target is shown as a mapping from a two dimensional factor-space. As can

be seen all optimum points lie on a circle in factor space. All these points are equivalent with respect to desirability. *There is no unique optimum.* This shows the need of investigating the structure of possible optima of DIs in later work. Insights in this field can lead to specialised search strategies for DI optimisation, therefore reducing the computing power that is needed in this process.

# 8 Aspects of modelling desirability indices

A prerequisite for applying any of the techniques given above is the estimation of individual models $\hat{f}_i$ for the targets $Y_i, i = 1, \ldots, Z$. These $\hat{f}_i$ are inserted into the individual DFs to calculate the DI for every setting $\mathbf{X}$ needed in the optimisation step.

## 8.1 Direct modelling

If the DIs can be estimated directly, an important decrease in computational complexity for solving the MCO can be expected. It is hoped, that not each of the $X_i$ which appear in any of the models for different targets will be important in a model for the DI, thus reducing the dimension of the relevant factor space. Nevertheless for both types of DIs mentioned the possible models come out to be very complicated, even if linear models for the $Y_i$ are assumed. As an example two targets $Y_i$, each a linear function of two factors $X_{1,2}$, an intercept $\alpha_i$ and an error $\epsilon_i$, $i = 1, 2$, no interactions are used.

In the Harrington case polynomials of high degree appear in the exponent of the DI:

$$\begin{aligned} D^H &= \sqrt{d_1^H(Y_1) \cdot d_2^H(Y_2)} \\ &= (\exp(-|(\alpha_1 + \gamma_{1,1}X_1 + \gamma_{1,2}X_2 + \epsilon_1)|^{n_1} \\ &\qquad -|(\alpha_2 + \gamma_{2,1}X_2 + \gamma_{2,2}X_2 + \epsilon_2)|^{n_2}))^{0.5} \end{aligned}$$

For the sake of simplicity it is assumed here that none of the two targets needs a further transformation, that is $Y_i = Y_i', \quad i = 1, 2$, which implies LSL=-1 and USL=1 for the Harrington case. Now it is clear that a model for $-\log(D^H)$ representing "the truth" has to be of order $\max_i(order(d_i^H(Y_i)^{n_i}))$ in the most simple case. Furthermore the error terms will interact with the effects if $n_i \neq 1$.

Under the same assumptions the Derringer/Suich DI leads to models containing all interactions of the main effects relevant for any of the $Y_i, i = 1, 2, \ldots, Z$.

$$\begin{aligned} D^{DS} &= \sqrt{d_1^{DS}(Y_1) \cdot d_2^{DS}(Y_2)} \\ &= ((\alpha_1 + \gamma_{1,1}X_1 + \gamma_{1,2}X_2 + \epsilon_1) \cdot \\ &\quad (\alpha_2 + \gamma_{2,1}X_1 + \gamma_{2,2}X_2 + \epsilon_2))^{0.5} \end{aligned}$$

In this latter case the most simple assumptions were made, too. Only one branch of the DF is considered for both targets, the exponents were set to unity. A model for $(D^{DS})^2$ would have to be of order $\prod_i(order(d_i))$. The usage of non-integer importance parameters $\beta_l$ and $\beta_r$ results in models of non-integer degrees.
As a bottom line direct modelling of DIs does not seem to be an option.

## 8.2 Sensitivity to modelling errors: a simulation case study

Another aspect, which has not found the attention it deserves, is the uncertainty in estimating the optimum factor setting $X_{opt}$. To examine the effect of estimating the underlying functions $f_i,\ i = 1, 2, \ldots, Z,$ with some error a simulation case study has been performed.
In their paper Derringer and Suich use a set of chemical data to apply their DFs to. They have four targets $Y_1$ to $Y_4$ and three controllable variables $X_1$ to $X_3$. The data were generated using a central-composite design with 20 experiments, to be able to fit a second-order model $\hat{f}_i$ including all interactions to each of the four targets. Using their data they got the following models for the targets $Y_i$:

$$\begin{aligned} \hat{f}_1 &= 139.1 + 16.5X_1 + 17.9X_2 + 10.9X_3 - 4.0X_1^2 - 3.5X_2^2 - 1.6X_3^2 \\ &\quad +5.1X_1X_2 + 7.1X_1X_3 + 7.9X_2X_3; \quad sd = 5.6, \\ \hat{f}_2 &= 1261.1 + 268.2X_1 + 246.5X_2 + 139.5X_3 - 83.6X_1^2 - 124.8X_2^2 \\ &\quad +199.2X_3^2 + 69.4X_1X_2 + 94.1X_1X_3 + 104.4X_2X3; \quad sd = 328.7, \\ \hat{f}_3 &= 400.4 - 99.7X_1 - 31.4X_2 - 73.9X_3 + 7.9X_1^2 + 17.3X_2^2 + 0.4X_3^2 \\ &\quad +8.8X_1X_2 + 6.3X_1X_3 + 1.3X_2X_3; \quad sd = 20.6, \\ \hat{f}_4 &= 68.9 - 1.4X_1 + 4.3X_2 + 1.6X_3 + 1.6X_1^2 + 0.1X_2^2 - 0.3X_3^2 \\ &\quad -1.6X_1X_2 + 0.1X_1X_3 - 0.3X_2X_3; \quad sd = 1.27. \end{aligned}$$

Targets $Y_1$ and $Y_2$ are of the type "the more the better". Targets $Y_3$ and $Y_4$ are real target value problems. All DFs in the original paper were linear. A numerical optimisation results in an optimum factor setting $X_{opt}^{DS} = (-0.05, 0.145, -0.868)$
For a simulation these four estimated functions now worked as "known world". The experimental design Derringer and Suich performed was 1000 times repeated in simulation, using the fitted models $\hat{f}_i$ as known $f_i, i = 1, 2, 3, 4$, and adding random errors according to the estimated standard deviations. Optimally the 1000 simulated optima $\hat{X_{Opt}}$ should scatter around the "true" $X_{Opt}^{DS}$, giving overall desirabilities near the optimum desirability. *Unfortunately this is not the case for the Derringer/Suich data!*
The result of the simulations can be assessed best looking at Figure 5. We find the expected scatter around the "true" optimum $X_{Opt}^{DS}$ in each of the two-dimensional projections, but circa 5% of the $\hat{X_{Opt}}$ show up far away from $X_{Opt}$. A closer look reveals, that many of these points predicted as optima give an overall desirability of 0, if evaluated with the known, true desirabilities! These points are marked with crosses in Figure 5.
The cause for this phenomenon lies in the big standard deviation of target $Y_2$. This results in poorly estimated $\hat{f}_2$ in simulations and from time to time this results in a qualitatively different response surface for the DI, giving nonsense optima.

# 9 Conclusion

Desirability is a concept for solving MCO based on rescaling and weighting individual targets. For assigning the weights a strong collaboration between the statistician and the expert is a must!
Direct modelling of the DI does not seem to be useful for reducing dimension of relevant factor space.
The unsatisfying results in the simulation study will lead to research concerning the interaction of estimating relations $f_i$, setting lower and upper bounds for poorly fitted targets and the sensitivity of the prognoses of $X_{Opt}$. It is hoped to find hints how DIs can be turned into self-diagnostic tools, identifying critical targets or factors and those reducing their sensitivity.

# References


[1] E. del Castillo, D.C. Montgomery, and D. R. McCarville. Modified desirability functions for multiple response optimization. *Journal of Quality Technology*, 28(3):337–344, July 1996.

[2] G. Derringer and R. Suich. Simultaneous optimization of several response variables. *Journal of Quality Technology*, 12(4):214–219, 1980.

[3] E.C. Harrington Jr. The desirability function. *Industrial Quaily Control*, 21(10):494–498, 1965.

[4] M.M.W.B. Hendriks, J. H. de Boer, A.K. Smilde, and D.A. Doornbos. Multicriteria decision making. *Chemometrics and intelligent laboratory sytems*, (16):175–191, 1992.

[5] W. Seewald and C. Weihs. *STAVEX User Manual.* CIBA-GEIGY Ltd., 4.100 edition, 1995.

Figure 5: Simulated optima for Derringer/Suich data as input model. The crossing dashed lines give the analytically correct optimum.