Proceedings Paper:

Duro, J.A. and Saxena, D.K. (2017) Timing the decision support for real-world many-objective optimization problems. In: Trautmann, H., Rudolph, G., Klamroth, K., Schütze, O., Wiecek, M., Jin, Y. and Grimme, C., (eds.) Evolutionary Multi-Criterion Optimization : 9th International Conference, EMO 2017, Proceedings. 9th International Conference on Evolutionary Multi-Criterion Optimization, 19-22 Mar 2017, Münster, Germany. Lecture Notes in Computer Science (10173). Springer International Publishing , pp. 191-205. ISBN 9783319541563

https://doi.org/10.1007/978-3-319-54157-0_14

Timing the Decision Support for Real-World Many-Objective Optimization Problems

João A. Duro[1] and Dhish Kumar Saxena[2]

[1] Automatic Control and Systems Engineering Department, The University of Sheffield, Sheffield, UK, j.a.duro@sheffield.ac.uk
[2] Department of Mechanical & Industrial Engineering, Indian Institute of Technology, Roorkee, India, dhishfme@iitr.ac.in

Abstract. Lately, there is growing emphasis on improving the scalability of *multi-objective evolutionary algorithms* (MOEAs) so that many-objective problems (characterized by more than three objectives) can be effectively dealt with. Alternatively, the utility of integrating *decision maker's* (DM's) preferences into the optimization process so as to target some most preferred solutions by the DM (instead of the whole Pareto-optimal front), is also being increasingly recognized. The authors here, have earlier argued that despite the promises in the latter approach, its practical utility may be impaired by the lack of—*objectivity*, *repeatability*, *consistency*, and *coherence* in the DM's preferences. To counter this, the authors have also earlier proposed a machine learning based decision support framework to reveal the preference-structure of objectives. Notably, the revealed preference-structure may be sensitive to the timing of application of this framework along an MOEA run. In this paper the authors counter this limitation, by integrating a termination criterion with an MOEA run, towards determining the appropriate timing for application of the machine learning based framework. Results based on three real-world many-objective problems considered in this paper, highlight the utility of the proposed integration towards an *objective*, *repeatable*, *consistent*, and *coherent* decision support for many-objective problems.

1 Introduction

Multi-objective evolutionary algorithms (MOEAs) have been well known to approximate the true *Pareto-optimal front* (POF) for a given multi-objective optimization problem, without emphasizing one objective over the other [1]. Notably, unlike the case of two- and three-objective problems, the performance of most existing MOEAs deteriorates as the number of objectives (M) grow beyond three [2]. This perhaps explains as to why optimization problems comprising of four or more objectives are distinctively referred to as *many-objective problems* (MaOPs) and have been receiving a lot of attention recently. The challenges associated with MaOPs relate to: (a) the nature of high-dimensional problems, and (b) the manner in which the existing MOEAs discriminate between better and worse solutions (*selection* operation). In terms of the former, visualization of a high-dimensional search space is difficult, and a good approximation (complete convergence and full coverage) of a high-dimensional POF calls for an (impractical) exponential increase in population size with a linear increase in M [3]. In terms of the latter, the *selection* operation either becomes ineffective or computationally too demanding. For instance: (a) the most commonly used *primary selection*, as in NSGA-II [4], is based on Pareto-dominance which fails to induce an effective partial order on the solutions as the number of objectives increases [5], (b) there is huge computational cost involved in dealing with a large number of weight vectors in decomposition based MOEAs, such as MOEA/D [6], and (c) the indicator based MOEAs, such as HypE [7] become impractical since indicators like hypervolume are computationally too demanding.

Acknowledging the poor scalability of most existing MOEAs with the number of objectives, an emergent strategy is to target only a handful of optimal or near-optimal solutions that are most preferred by the *decision maker* (DM). The engagement of a DM with the optimization process has led to what is referred to as *multiple criteria decision making* (MCDM) based MOEAs [8]. The authors in [5] have argued that despite their promise, the utility of MCDM based MOEAs may be impaired by the lack of–*objectivity*, *repeatability*, *consistency*, and *coherence* in the DMs' preferences. Towards countering these limitations, the authors have proposed a framework for machine learning based decision support:

1. expressed through *revelation* of the preference–structure of different objectives embedded in the problem model,
2. which can potentially aid the DMs to induce a preference–order over the solutions (guided by the preference–order of the objectives) with *objectivity*, *repeatability*, *consistency*, and *coherence*.

This paper distinguishes between: (δ-I) the *capability of a decision support* framework to *reveal* the preference–structure of different objectives in a *given* input solution set (non-dominated solutions obtained from an MOEA), and (δ-II) the *capability to determine the timing of decision support*, implying, the capability to determine the number of generations along an MOEA run at/after which, the corresponding non-dominated solution set could be treated as the most *appropriate* input solution set for application of the machine learning based framework for revelation of the objectives' preference–structure. While the former has been demonstrated in [5], this paper focuses on the latter—a more fundamental aspect of *when to time the decision support*. The criticality of the latter aspect can be gauged from the fact that it is analogous to the fundamental and largely unaddressed question (until [9]) of when to terminate an MOEA in the absence of a termination criterion that is *robust* in terms of its: (a) *generality*, implying that it does not require an a priori knowledge of the POF, and neither depends on MOEA-specific operators, nor on the MOEA-related performance indicators, (b) *on-the-fly* implementation, and (c) *computational efficiency* enabling efficient scalability with the number of objectives. In the absence of [9], the capability of the framework proposed in [5] was demonstrated on input solution sets corresponding to *a priori* fixed number of generations. However, despite its novelty and theoretical contribution, the practical utility of [5] stands implicitly impaired by the fact that if the *a priori* fixed number of generations are not sufficient to ensure that the input solution set corresponds to a *stabilized* MOEA population, the *revealed* objectives' preference–structure and solutions' preference–order can not be treated as accurate, robust and efficient representatives of the problem–structure. In that, if the *a priori* fixed number of generations are:

A. too low: the MOEA population may not stabilize and the *revealed* problem–structure may itself change if the underlying MOEA was allowed to run for more generations, rendering the offered decision support futile,
B. too high: the MOEA population may stabilize far earlier, implying wastage of computational resources despite no variation in *revealed* problem–structure.

In view of the above, this paper aims to bridge the gap between *capability of a decision support* framework (δ-I) and *capability to determine the timing of decision support* (δ-II), and achieves this by integrating the capabilities developed in [5] and [9]. The utility of this work in terms of the resulting *robust* decision support has been highlighted through three real-world instances of many-objective optimization problems. In that, it is demonstrated: (a) how an ad hoc *a priori* fixation of the timing of application of the framework for decision support [5] may lead to misleading objectives' preference–structure, and (ii) how its integration with an entropy based MOEA–termination criterion can help overcome this challenge, leading up to a *robust* decision support.

The structure of the remaining paper is as follows. The decision-support framework [5] and the MOEA-termination algorithm [9] are briefly described in Section 2. This is followed by a brief description of the considered real-world MaOPs in Section 3. The results and associated discussions are presented in Section 4, while the paper concludes with Section 5.

2 Methodology

It has been highlighted above that aiming at an accurate, robust and efficient decision support, this paper integrates a framework for machine learning based decision support [5] and an MOEA-termination algorithm [9]. For completeness, each of these are summarized below.

2.1 Framework for Machine Learning based Decision support

The framework for machine learning based decision support proposed in [5], *reveals* the preference–structure of different objectives in terms of:

1. an *essential* objective set, implying a smallest set of conflicting objectives which can generate the same POF as the original problem,
2. ranking of *essential* objectives which further paves the way for revelation of: (i) the smallest objective sets corresponding to pre-specified errors, and (ii) objective sets of pre-specified sizes that correspond to minimum error.

This framework is based on the premise that: (i) most existing MOEAs provide a poor POF-approximation for MaOPs, implying that the dominance relations characterizing the obtained solutions may be different from those characterizing the true POF, and (ii) the dimensionality of the POF need not be the same as the number of objectives M. The problem of *learning* objectives' preference–structure on the true POF, from a solution set that may not represent the true POF, is posed as a machine learning problem. In that, given a set of non-dominated solutions, the objectives' preference–structure is *learnt* by application of PCA (Principal Component Analysis: relying on the correlation matrix R) or MVU (Maximum Variance Unfolding: relying on the kernel matrix K), aided by:

1. an interpretation of objectives as conflicting or non-conflicting based on their relationship along the eigenvectors of R or K,
2. a dynamic interpretation of the correlation between the objectives.

Depending on the choice of R or K, the framework allows for two variants, namely L-PCA based decision support and NL-MVU-PCA based decision support. The two can be distinguished in terms of the former's inability to account for nonlinearity, unlike the latter. While the details of the framework can be found in [5], its building blocks are summarized below.

Computation of Correlation and kernel Matrices Let a non-dominated solution set be obtained by running an MOEA with the initial objective set $\mathcal{F}_0 = \{f_1, \ldots, f_M\}$ and a population size of N. Let the objective vector in the non-dominated set, corresponding to the i^{th} objective (f_i) be denoted by $\dot{f}_i \in \mathbb{R}^N$ and its mean and standard deviation by $\mu_{\dot{f}_i}$ and $\sigma_{\dot{f}_i}$, respectively. Furthermore, let $\ddot{f}_i = (\dot{f}_i - \mu_{\dot{f}_i})/\sigma_{\dot{f}_i}$. Then, the input data X and the correlation matrix R can be composed as $X_{M\times N} = [\ddot{f}_1 \; \ddot{f}_2 \ldots \ddot{f}_M]^T$ and $R_{M\times M} = \frac{1}{M} XX^T$. Furthermore, the kernel matrix K can be *learnt* from a semidefinite programming problem, presented in Equation 1.

$$
\begin{aligned}
&\text{Maximize } trace(K) = \textstyle\sum_{ij} \frac{(K_{ii} - 2K_{ij} + K_{jj})}{2M} \\
&\text{subject to the following constraints :} \\
&(a) K_{ii} - 2K_{ij} + K_{jj} = R_{ii} - 2R_{ij} + R_{jj}, \forall\ \eta_{ij} = 1 \\
&(b) \textstyle\sum_{ij} K_{ij} = 0 \\
&(c) K \text{ is positive-semidefinite,} \\
&\text{where : } R_{ij} \text{ is the } (i,j)^{th} \text{ element of the correlation matrix } R \\
&\qquad\qquad \eta_{ij} = \begin{cases} 1, & \text{if } \dot{f}_j \text{ is among the } q\text{-nearest neighbor of } \dot{f}_i \\ 0, & \text{otherwise} \end{cases}
\end{aligned}
\tag{1}
$$

Objective reduction based on Eigenvalue Analysis This step aims to: (i) interpret the objectives as conflicting or non-conflicting based on their relationship along the significant eigenvectors of R or K (depending on the choice of L-PCA based or NL-MVU-PCA based decision support), and (ii) retain only the conflicting objectives. Towards it, obtain the eigenvalues of R or K, and sort these in descending order of magnitude as $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M$. Let the corresponding eigenvectors be given by V_1, V_2, ..., V_M, the contribution of the i^{th} objective towards V_j be given by f_{ij}, and the normalized eigenvalues be given by $e_j = \lambda_i / \sum_{j=1}^{M} \lambda_i$ (implying $\sum_{j=1}^{M} e_j = 1$). Subsequently, determine the number of *significant* eigenvectors as the smallest number (N_v) such that $\sum_{j=1}^{N_v} e_j \geq \theta$, where the variance threshold $\theta \in [0, 1]$ is an algorithm parameter prescribed to be set to $\theta = 0.997$. Then, the set of objectives ($\mathcal{F}_e$) based on conflict along significant eigenvectors can be composed by picking for each significant V_js: the objective with the highest contribution (in terms of magnitude) along with all other opposite sign objectives. As an exception, if all objectives have the same-sign contribution along a particular V_j, then the objectives with top two contributions by magnitude are selected. Notably $\mathcal{F}_e \subseteq \mathcal{F}_0$.

Objective reduction based on Set-based Correlation This step aims to: (i) identify the subsets of correlated objectives within $\mathcal{F}_e$, and (ii) retain from each such correlated subset, only the most significant objective while discarding the rest. Towards it, for each $f_i \in \mathcal{F}_e$, constitute a subset $\mathcal{S}_i$ with correlated objectives $f_{j|j \neq i} \in \mathcal{F}_e$ based on the following:

$$f_j \in \mathcal{S}_i : \begin{cases} sign(R_{ik}) = sign(R_{jk}),\ \forall\ k = 1, 2, \ldots, M, \\ R_{ij} \geq T_{cor} = 1.0 - e_1(1.0 - M_{2\sigma}/M), \end{cases}$$

where T_{cor} is the correlation threshold and $M_{2\sigma}$ is the smallest number, such that $\sum_{j=1}^{M_{2\sigma}} e_j \geq 0.954$. Subsequently, for each objective in $\mathcal{S}_i$, assign a score sc_i given by:

$$sc_i = \sum_{j=1}^{N_v} e_j |f_{ij}|$$

following which, the objective with the highest sc_i is retained while the other objectives in $\mathcal{S}_i$ are eliminated. This step facilitating further objective reduction leads to an *essential* objective set $\mathcal{F}_s$, where $\mathcal{F}_s \subseteq \mathcal{F}_e \subseteq \mathcal{F}_0$.

Preference-ranking of all objectives The preference-ranking of objectives is guided by the errors to be incurred on elimination of objectives. An error associated with elimination of a particular objective corresponds to the variance that is left unaccounted if that objective is eliminated alone, and can be computed as in Equation 2, depending on whether it belongs to an *essential* or *redundant* objective set.

$$\left.\begin{array}{ll} \mathcal{E}_i & = c_i^M \text{ for } f_i \in \mathcal{F}_s \\ \mathcal{E}_i & = c_i^M (1.0 - \max_{j \in \mathcal{F}_s} \{\delta_{i\,j}.R_{ij}\}) \text{ for } f_i \in \mathcal{F}_{redn} \equiv \mathcal{F}_0 \setminus \mathcal{F}_s \\ \text{where:} & \\ c_i^M & = \sum_{k=1}^{M} e_k f_{ik}^2 \\ \delta_{i\,j} & = \begin{cases} 1, & \text{if } f_i \text{ and } f_j \text{ are correlated} \\ 0, & \text{otherwise} \end{cases} \\ R_{ij} & = \text{Correlation between} f_i \text{ and} f_j \end{array}\right\} \quad (2)$$

Finally, the preference-weight for each objective is given by Equation 3 and the preference-ranking of all objectives is established by the sorted $\mathcal{E}_i^n$s. For instance, let u and v be two objectives such that $\mathcal{E}_u^n \gg \mathcal{E}_v^n$, implying that the error incurred by eliminating u is far greater than the error incurred if objective v were eliminated. In other words, for higher accuracy, the objective u needs to be preferred over v (or the solutions which are better in u need to be preferred over those which are better in v).

$$w_i = \mathcal{E}_i^n = \mathcal{E}_i / \sum_{j=1}^{M} \mathcal{E}_j \quad \text{(ensuring that } w_i \geq 0 \text{ and } \sum_{i=1}^{M} w_i = 1\text{)} \qquad (3)$$

2.2 An Entropy based MOEA-termination Algorithm

The MOEA-termination algorithm proposed in [9] is based on the premise that it is prudent to terminate an MOEA if the MOEA population in successive generations does not undergo significant changes. Towards implementation of this principle:

1. a *dissimilarity* measure across two successive MOEA generations, based on information theory concepts of *entropy* and *relative entropy* was proposed in [9],
2. the multidimensional histogram method (Section IV in [9]) which relies on partitioning each dimension of the M-dimensional objective space into a fixed number of intervals (n_b), facilitates the computation of a *dissimilarity* measure (Section V in [9]),
3. the conformance of the mean and standard deviation of the *dissimilarity* measure up to a pre-specified accuracy level (n_p:number of decimal places), across a pre-specified number of successive MOEA generations (n_s), was treated as a termination criterion (Section VI in [9]).
4. just when the termination criterion is satisfied, the MOEA run is terminated and the number of MOEA generations up to that point are denoted by N_{gt}.

The resulting termination algorithm helps identify *on-the-fly* the number of generations beyond which an MOEA *stabilizes*, implying that either a good POF–approximation has been obtained, or that it can not be obtained due to the stagnation of the MOEA in the search space (excluding the POF). The fact that in either case, no further improvement in the POF–approximation can be obtained despite additional computational expense, provides the rationale for MOEA-termination. As highlighted earlier, the hallmark of the proposed algorithm (besides its *on-the-fly* implementation), lies in *generality* and *computational efficiency* enabling efficient scalability with the number of objectives.

3 Real-world Many-objective Optimization Problems

A brief description of the real-world instances of MaOPs considered is as follows.

3.1 Machining problem

The authors in [10] have investigated the multi-criteria optimization of machining operations, when applied to a 390 die cast aluminum alloy with VC-3 carbide cutting tools. The optimization problem consists of four objectives, three decision variables and three inequality constraints. The decision variables relate to the cutting speed (x_1), feed rate (x_2), and the depth of cut (x_3). The related constraints and objective functions are, as below.

$$\left.\begin{array}{rl}\text{Minimize} & f_1(\mathbf{x}) \equiv \text{Surface roughness},\\ \text{Maximize} & f_2(\mathbf{x}) \equiv \text{Surface integrity},\\ \text{Maximize} & f_3(\mathbf{x}) \equiv \text{Tool life},\\ \text{Maximize} & f_4(\mathbf{x}) \equiv \text{Metal removal rate},\\ & \text{subject to:}\\ & g_1(\mathbf{x}) \equiv \text{Upper limit on surface roughness}\\ & g_2(\mathbf{x}) \equiv \text{Lower limit on surface integrity}\\ & g_3(\mathbf{x}) \equiv \text{Lower limit on tool life.}\end{array}\right\} \tag{4}$$

3.2 Storm drainage system problem

Water resources management is a discipline that consists of planning, distributing, developing, and managing the optimum use of water resources. The example here deals with the optimal planning for a storm-drainage system located in a urban area as described in [11]. The decision variables relate to the local detention storage capacity (x_1), maximum treatment rate (x_2), and the maximum allowable overflow rate (x_3). The related constraints and objective functions are, as below.

$$\left.\begin{array}{l}\underset{(x_1,x_2,x_3)}{\text{Minimise:}}\\ \quad f_1(\mathbf{x}) \equiv \text{Drainage network cost},\\ \quad f_2(\mathbf{x}) \equiv \text{Storage facility cost},\\ \quad f_3(\mathbf{x}) \equiv \text{Treatment facility cost},\\ \quad f_4(\mathbf{x}) \equiv \text{Expected flood damage cost},\\ \quad f_5(\mathbf{x}) \equiv \text{Expected economic loss due to flood},\\ \text{subject to:}\\ \quad g_1(\mathbf{x}) \equiv \text{Average no. of floods/year},\\ \quad g_2(\mathbf{x}) \equiv \text{Probability of flood depth exceeding 1 basin-inch},\\ \quad g_3(\mathbf{x}) \equiv \text{Average no. of pounds/year of suspended solids},\\ \quad g_4(\mathbf{x}) \equiv \text{Average no. of pounds/year of settable solids},\\ \quad g_5(\mathbf{x}) \equiv \text{Average no. of pounds/year of } BOD,\\ \quad g_6(\mathbf{x}) \equiv \text{Average no. of pounds/year of } N,\\ \quad g_7(\mathbf{x}) \equiv \text{Average no. of pounds/year of } P0_4.\end{array}\right\} \tag{5}$$

3.3 Work roll cooling design problem

In metalworking, a rolling operation is the process of shaping a metal strip by reducing its thickness and creating a uniform surface. This is achieved by passing the metal between two rollers that are driven at equal peripheral speed in opposite directions. During this process, heat is transferred from the metal strip to the rolls and in case it becomes too excessive it can lead to the formation of roll cracks or to any other types of damage. The decision variables [12] relate to the roll/stock contact HTC (x_1), stock temperature (x_2), roll/stock contact length (x_3), cooling HTC (x_4), roll speed (x_5), roll temperature (x_6), and delay time (x_7). The related constraints and objective functions are, as below.

$$\left.\begin{array}{l}\underset{(x_1,\ldots,x_7)}{\text{Minimise:}}\\ \quad f_1(\mathbf{x}) \equiv \text{Change in temperature at roll surface,}\\ \quad f_2(\mathbf{x}) \equiv \text{Radial stress at the roll surface,}\\ \quad f_3(\mathbf{x}) \equiv \text{Change in temperature at 9mm depth,}\\ \quad f_4(\mathbf{x}) \equiv \text{Radial stress at 9mm depth,}\\ \quad f_5(\mathbf{x}) \equiv \text{Change in temperature at 15mm depth,}\\ \quad f_6(\mathbf{x}) \equiv \text{Radial stress at 15mm depth.}\end{array}\right\} \quad (6)$$

4 Experimental results

This section demonstrates the application of the decision support framework to three real-world problems. NSGA-II [4] has been used as the underlying MOEA, and the framework is applied to two populations that are chosen at different generations during the NSGA-II run. The first corresponds to the initial population of solutions ($N_g = 1$), while the second population is chosen by the MOEA-termination algorithm (N_{gt}). The analysis generated by the application of the framework to the two populations are compared and their differences are highlighted. The settings of the parameters involved, are as below:

1. the NSGA-II related parameters include a population size of 200; the probability of crossover and mutation as 0.9 and 0.1, respectively; the distribution index for crossover and mutation as 5 and 20, respectively; and the maximum number of generations as 10000,
2. the chosen parameters for the decision support framework are: $\theta = 0.997$; and the neighborhood size is given by $q = M - 1$ (Equation 1),
3. the chosen parameters for the MOEA-termination algorithm [9] are $n_b = 20$; $n_s = 20$; and $n_p = 2$.

Towards analysis of the results, in terms of a comparison between R_{ij} (strength of correlation between objectives f_i and f_j); f_{ij} (contribution of objective f_i along the principal component V_j); or the preference weight w_i for objective f_i, across different N_{gt}, the notion of *relative percentage difference* (RPD) is utilized. In that, if the two quantities being compared (say R_{ij} at two different N_{gt}) are denoted by α_1 and α_2, then RPD is given by Equation 7.

$$\text{RPD}(\alpha_1, \alpha_2) = |(\alpha_1 - \alpha_2)|/\max(|\alpha_1|, |\alpha_2|) \times 100, \quad (7)$$

(a) Machining problem

(b) Storm drainage system problem

(c) Work roll cooling design problem

Fig. 1: Preference-weights for the objectives captured along the optimization. The dashed vertical line corresponds to the instant where stability is detected by the termination criterion. The results correspond to one NSGA-II run.

It may be noted that in the tabular results presented ahead, the numbers in the row labeled as RPD(%) denote the RPD(α_1, α_2) values computed by treating the two quantities in the corresponding column above as α_1 and α_2.

Notably, this paper: (i) is based on the premise that the revealed preference-structure may be sensitive to the timing of application of the decision support framework, and (ii) aimed to determine the *appropriate timing* for application of the framework, such that the revealed preference-structure in neither misleading nor comes at an avoidable computational cost. While problem-specific results are discussed in greater details in the following sections, Figure 1 which presents a snapshot of results for all the problems, not only validates the basic premise of this paper, but also demonstrates that the aim has been realized. In that, regardless of the problem, the following can be observed:

1. there is significant variation in the preference-weight for each objective up to specific number of MOEA generations beyond which it largely stabilizes,
2. in the absence of integration of the decision support framework and MOEA-termination algorithm, when the suitable timing may not be otherwise known, it is highly probable that the *a priori* fixed number of MOEA generations may be such that the preference–structure is revealed either when the preference weights for some/all objectives are undergoing variation or long after they have stabilized. While in the former case, the preference–structure may be misleading, the latter case is marked by avoidable computational wastage.
3. integration of MOEA-termination algorithm has indeed revealed an *appropriate* timing of application of the decision support framework, since the dashed vertical line representing the prescribed timing in each plot is located only just after the preference-weights have stabilized and much before an instance of computational wastage can be cited.

4.1 Machining problem

For this problem, while the MOEA-termination algorithm infers $N_{gt} = 308$, the corresponding decision support revelations (Table 1) are also compared to those corresponding to *a priori* fixed $N_g = 1$. In that:

1. There is a change in the correlation structure from $N_{gt} = 1$ to $N_{gt} = 308$, since the correlation sign for R_{24} reversed from positive to negative. There are also significant changes to the correlation strengths between some objective pairs. For instance, the change in correlation for R_{13}, and also for R_{23}, is approximately 70%. The analysis of the decision support framework indicates that all objectives are uncorrelated since no two pair of objectives share the same correlation signs with the remaining objectives. One example is f_1 that is positively correlated with f_2 and f_3 (i.e., $R_{12} > 0$ and $R_{13} > 0$), however, since the correlation between f_2 and f_3 is negative (i.e., $R_{23} < 0$), f_1 and f_2 are not interpreted as correlated.
2. There are significant changes to the principal components of the problem. In that: (i) at $N_g = 1$, the first principal component accounts for 74.7079% of the variance (i.e., $e_1 = 0.747079$); and, (ii) at $N_{gt} = 308$, this number rises to 93.5543% (i.e., $e_1 = 0.935543$). It is therefore expected for the preference-weights that are determined at $N_{gt} = 308$, to rely mostly on the first principal component. A high RPD is reported when comparing the elements of the first principal component, that is, almost 200% in some cases. This indicates that the distribution of solutions, across the search space, have experienced a change in direction while approaching stability. The value of T_{cor} suffers a reduction, with an RPD of 15.042%, which is mostly attributed to the increase in e_1.
3. Since no objectives can be interpreted as being correlated due to their correlation signs, the essential objective set is $\mathcal{F}_s = \{f_1, f_2, f_3, f_4\}$. This remains unaltered during the optimization process.

Table 1: Key highlights for the machining problem.

Gens.	Correlation structure						
	R_{12}	R_{13}	R_{14}	R_{23}	R_{24}	R_{34}	T_{cor}
$N_g = 1$	0.512585	0.222373	-0.786626	-0.715817	0.079025	-0.750730	0.626461
$N_{gt} = 308$	0.498700	0.734629	-0.913131	-0.208278	-0.145054	-0.936288	0.532228
RPD(%)	2.7088	69.730	13.854	70.903	154.48	19.818	15.042

Gens.	First principal component				Second principal component			
	f_{11}	f_{21}	f_{31}	f_{41}	f_{12}	f_{22}	f_{32}	f_{42}
$N_g = 1$	0.5581	-0.0666	0.2849	-0.7765	-0.4955	-0.2911	0.8178	-0.0312
$N_{gt} = 308$	-0.5404	0.0359	-0.2860	0.7905	-0.5558	-0.1740	0.8091	-0.0793
RPD(%)	196.82	153.84	199.61	198.22	10.839	40.223	1.0666	60.694

Gens.	e_1	e_2	Preference-weights				F_s
			w_1	w_2	w_3	w_4	
$N_g = 1$	0.747079	0.249107	0.2946281	0.02694043	0.2272319	0.4511996	$\{f_1, f_2, f_3, f_4\}$
$N_{gt} = 308$	0.935543	0.062604	0.2927858	0.00442904	0.1175300	0.5852552	$\{f_1, f_2, f_3, f_4\}$
RPD(%)	20.145	74.869	0.62530	83.560	48.278	22.905	—

4. There are some significant changes to the preference-weights. In that, a reduction that corresponds to an RPD of 83.569% and 48.278% is reported, respectively, for w_2 and w_3. While an increase that corresponds to an RPD of 22.905% is reported for w_4. The preference-weight w_1 remains almost unaltered. The change in the preference-weights is mostly attributed to the changes experienced by the principal components since all objectives are interpreted as essential. The preference order between the objectives remains unaltered throughout the optimization process and is given by $w_4 > w_1 > w_3 > w_2$.

The framework analysis for the Machining problem has highlighted some important differences when comparing $N_g = 1$ with a population obtained at a stable instant. It is worthy to note the differences in the correlations between the objectives, implying that the correlation structure changes during the optimization process, as solutions approach stability. This could have implications for the problem analysis when validating the relationships between the objectives, in light of their physical meanings. Another important aspect relates to the distribution of the problem variance along the principal components. In that, it has been shown that this can have a significant influence on the accuracy of the preference-weights.

4.2 Storm drainage system problem

For this problem, while the MOEA-termination algorithm infers $N_{gt} = 195$, the corresponding decision support revelations (Table 2) are also compared to those corresponding to *a priori* fixed $N_g = 1$. In that:

1. There are changes to the correlation structure since some correlation signs have been reversed. This is the case for R_{12}, R_{23} and R_{24}. There are significant changes to the strength of correlation between some objectives. This is noteworthy for objective f_2 where the correlations R_{12}, R_{23} and R_{24}, all show an RPD that is higher than 100%. Another significant change is between f_1 and the other objectives where R_{13}, R_{14} and R_{15}, correspond to an RPD of 30.580%, 59.800% and 94.495%, respectively.
2. The eigenvalues analysis shows that there are no significant changes to the principal components. In that, the variance is accounted for mostly by the first principal component (i.e., $e_1 \approx 0.98$), and the RPD is only 0.228%. As a result, T_{cor} is relatively low compared to most correlation strengths between the objectives, which implies that this is a highly redundant problem. The elements of the principal components have a RPD that ranges between 2.6376% and 31.307%. This suggests that the majority of the solutions have kept the same direction in the search space while approaching stability.
3. There are changes to the identified correlated sets. In that: (i) at $N_g = 1$ the sets are $\mathcal{S}_1 = \mathcal{S}_3 = \{f_1, f_3\}$, $\mathcal{S}_2 = \{f_2\}$ and $\mathcal{S}_4 = \mathcal{S}_5 = \{f_4, f_5\}$; and, (ii) at $N_{gt} = 195$ the sets are $\mathcal{S}_1 = \mathcal{S}_3 = \{f_1, f_3\}$, $\mathcal{S}_2 = \{f_2\}$ and $\mathcal{S}_4 = \{f_4\}$ $\mathcal{S}_5 = \{f_5\}$. This is attributed to the changes in the correlation structure, where objectives f_4 and f_5 are no longer interpreted as being correlated at $N_{gt} = 195$. In that, note the changes to the correlation signs where R_{24} is now positive while R_{25} remains negative. This implies that f_4 and f_5 react differently to the presence of f_2, hence, they cannot be interpreted as correlated. This also means that the essential objective set has changed during the optimization process, since $\mathcal{F}_s = \{f_2, f_3, f_4\}$ at $N_g = 1$, while at $N_{gt} = 195$ $\mathcal{F}_s = \{f_2, f_3, f_4, f_5\}$.
4. Despite the changes to the correlated sets, there are no major changes to the preference-weights. In that, the RPD ranges from 3.5827% to 99.226%. The preference order between the objectives is kept the same during the optimization process, which is $w_4 > w_3 > w_2 > w_5 > w_1$.

The above shows a situation where there are changes to the analysis of the decision support framework, first to the correlation structure of the problem, and then to the essential objective set. The fact that one objective is erroneously interpreted as being redundant can have severe consequences for the problem decision making task. This further highlights the importance of detecting stability during an optimization process, before the application of a decision support framework.

4.3 Work roll cooling design problem

For this problem, while the MOEA-termination algorithm infers $N_{gt} = 167$, the corresponding decision support revelations (Table 3) are also compared to those corresponding to *a priori* fixed $N_g = 1$. In that:

1. There are no changes to the correlation signs, implying that the correlation structure remains unaltered. Most objectives are highly correlated due to the

Table 2: Key highlights for the storm drainage system problem.

	Correlation structure									
Gens.	R_{12}	R_{13}	R_{14}	R_{15}	R_{23}	R_{24}	R_{25}	R_{34}	R_{35}	R_{45}
$N_g = 1$	0.0709	0.6916	-0.3657	-0.0234	0.0878	-0.0124	-0.5794	-0.8378	-0.3399	0.3908
$N_{gt} = 195$	-0.3035	0.9963	-0.9096	-0.4246	-0.2863	0.3563	-0.5420	-0.9187	-0.4492	0.4895
RPD(%)	123.36	30.580	59.800	94.495	130.67	103.49	6.4535	8.8141	24.310	20.165

	First principal component					Second principal component				
Gens.	f_{11}	f_{21}	f_{31}	f_{41}	f_{51}	f_{12}	f_{22}	f_{32}	f_{42}	f_{52}
$N_g = 1$	0.1586	0.1580	0.4012	-0.8743	0.1566	-0.3238	-0.3319	0.7994	0.1886	-0.3323
$N_{gt} = 195$	0.1254	0.1240	0.4786	-0.8513	0.1233	-0.3408	-0.3448	0.7556	0.2745	-0.3445
RPD(%)	20.945	21.505	16.181	2.6376	21.272	4.9939	3.7450	5.4787	31.307	3.5438

				Preference-weights					
Gens.	T_{cor}	e_1	e_2	w_1	w_2	w_3	w_4	w_5	$\mathcal{F}_s$
$N_g = 1$	0.2117	0.9853	0.0147	0.0084	0.0269	0.1728	0.7757	0.0162	$\{f_2, f_3, f_4\}$
$N_{gt} = 195$	0.2099	0.9876	0.0124	0.0001	0.0169	0.2373	0.7289	0.0168	$\{f_2, f_3, f_4, f_5\}$
RPD(%)	0.8520	0.228	15.346	99.226	37.160	27.176	6.0300	3.5827	—

reported high correlation values. The evolution of the correlation strength between pairs of objectives shows a low RPD which ranges between 0.5015% to 27.948%.

2. Most of the problem variance is accounted for by the first principal component with a tendency to increase as the population of solutions approaches stability. In that, e_1 increases from 0.9173 to 0.9898, which corresponds to a RPD of 7.3267%. This increase in e_1 is achieved at the cost of the second and third principal components since their values have decreased with a RPD of 87.748% and 86.528%, respectively. The analysis also shows that this is a highly redundant problem since the T_{cor} value is significantly smaller than the reported correlation values. As the solutions approach stability, the value of T_{cor} decreases from 0.3885 to 0.1752 which corresponds to a RPD of 54.905%.
3. With a low T_{cor} and a relatively high correlation strength values, the correlated sets are $\mathcal{S}_1 = \mathcal{S}_3 = \mathcal{S}_5 = \{f_1, f_3, f_5\}$ and $\mathcal{S}_2 = \mathcal{S}_4 = \mathcal{S}_6 = \{f_2, f_4, f_6\}$. The objectives f_1 and f_4 are identified as the essential objectives of the correlated sets, due to their contribution score towards the principal components. That is, although only shown for the first principal component it can be seen that, $|f_{11}| > |f_{31}| > |f_{51}|$ and $|f_{41}| > |f_{61}| > |f_{21}|$. This leads to the essential objective set given by $\mathcal{F}_s = \{f_1, f_4\}$.
4. The preference-weights for objectives in $\mathcal{F}_s$ have a low RPD when compared with the other preference-weights. That is, although the RPD for w_1 and w_4 is only 11.331% and 5.6665%, respectively, the other preference-weights show a RPD higher than 84%. This is mostly due to the changes in the correlation strength between redundant and essential objectives, where an increase in the correlation strength leads to a higher preferability for the essential objectives.

Table 3: Key highlights for the work roll cooling design problem.

Gens.	R_{12}	R_{13}	R_{14}	R_{15}	R_{16}	R_{23}	R_{24}	R_{25}	R_{26}	R_{34}	R_{35}
	Correlation structure										
$N_g = 1$	-0.7652	0.9566	-0.9529	0.9325	-0.9656	-0.7474	0.8688	-0.6885	0.7832	-0.9737	0.9939
$N_{gt} = 167$	-0.9719	0.9943	-0.9951	0.9899	-0.9954	-0.9656	0.9813	-0.9556	0.9705	-0.9968	0.9989
RPD(%)	21.263	3.8011	4.2367	5.8041	2.9936	22.605	11.461	27.948	19.296	2.3153	0.5015

Gens.	R_{36}	R_{45}	R_{46}	R_{56}	T_{cor}	f_{11}	f_{21}	f_{31}	f_{41}	f_{51}	f_{61}
	Correlation structure					First principal component					
$N_g = 1$	-0.9936	-0.9491	0.9862	-0.9833	0.3885	-0.4424	0.3675	-0.4048	0.4647	-0.3728	0.3879
$N_{gt} = 167$	-0.9988	-0.9929	0.9984	-0.9972	0.1752	-0.4564	0.3602	-0.3981	0.4665	-0.3639	0.3917
RPD(%)	0.5201	4.4183	1.2211	1.3912	54.905	3.0682	1.9849	1.6713	0.3909	2.3803	0.9793

Gens.	e_1	e_2	e_3	w_1	w_2	w_3	w_4	w_5	w_6	$\mathcal{F}_s$
				Preference-weights						
$N_g = 1$	0.9173	0.0778	0.0042	0.4285	0.0539	0.0154	0.4766	0.0204	0.0052	$\{f_1, f_4\}$
$N_{gt} = 167$	0.9898	0.0095	0.0006	0.4831	0.0059	0.0021	0.5052	0.0031	0.0006	$\{f_1, f_4\}$
RPD(%)	7.3267	87.748	86.528	11.311	89.068	86.541	5.6665	84.896	88.823	—

The preference order between the objectives remains unaltered throughout the optimization process and is given by $w_4 > w_1 > w_2 > w_5 > w_3 > w_6$.

The above situation corresponds to a highly redundant problem where the solutions are mostly spread across objectives f_1 and f_4. The analysis of Fig. 1c shows that the redundant objectives have a reduction in their preference-weights while the essential objectives have a slight increase. There is a high RPD for some preference-weights, such as w_2, w_3, w_5 and w_6, where the RPD is 89.068%, 86.541%, 84.896% and 88.823%, respectively. It is in such cases that the application of an MOEA-termination algorithm becomes significant.

5 Conclusion

This paper presents a framework for an *appropriately timed* and *accurate* decision support in the context of both multi- and many-objective optimization problems. The key features and contributions of this framework based on integration of machine learning based decision support and MOEA-termination algorithm have been highlighted in the context of three real-world problems. This paper is particularly significant in the context of many-objective optimization problems where the user or the decision maker may otherwise be clueless about the number of generations that may be sufficient for the corresponding MOEA population to be treated with some degree of confidence in terms of its conformance or departure from the true POF. Though for the sake of brevity, the *dissimilarity measure* leading up to the prescribed termination has not been shown, reference to it shall also indicate whether or not the *revealed* preference–structure could be treated as representative of the true POF. It is hoped that this integrated framework shall be instrumental in facilitating decision makers' preferences characterized by *objectivity*, *repeatability*, *consistency*, and *coherence* and shall mark a significant contribution towards the practical utility of MCDM based MOEAs.

References

1. C. A. C. Coello, G. B. Lamont, and D. A. V. Veldhuizen, *Evolutionary Algorithms for Solving Multi-Objective Problems*. Boston: Kluwer Academic Publishers, 2002.
2. R. Purshouse and P. Fleming, "Evolutionary many-objective optimisation: an exploratory analysis," in *Congress on Evolutionary Computation*, vol. 3, pp. 2066–2073, December 2003.
3. H. Ishibuchi, N. Tsukamoto, and Y. Nojima, "Evolutionary many-objective optimization: A short review," in *IEEE Congress on Evolutionary Computation*, pp. 2419–2426, June 2008.
4. K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A Fast Elitist Multi-Objective Genetic Algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2000.
5. J. A. Duro, D. K. Saxena, K. Deb, and Q. Zhang, "Machine Learning based Decision Support for Many-objective Optimization Problems," *Neurocomputing*, vol. 146, no. 0, pp. 30–47, 2014.
6. Q. Zhang and H. Li, "MOEA/D: A Multiobjective Evolutionary Algorithm Based on Decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 11, pp. 712–731, December 2007.
7. J. Bader and E. Zitzler, "HypE: An Algorithm for Fast Hypervolume-based Many-objective Optimization," *Evolutionary Computation*, vol. 19, no. 1, pp. 45–76, 2011.
8. R. C. Purshouse, K. Deb, M. M. Mansor, S. Mostaghim, and R. Wang, "A Review of Hybrid Evolutionary Multiple Criteria Decision Making Methods," in *IEEE Congress on Evolutionary Computation*, pp. 1147–1154, July 2014.
9. D. K. Saxena, A. Sinha, J. A. Duro, and Q. Zhang, "Entropy based Termination Criterion for Multiobjective Evolutionary Algorithms," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 4, pp. 485–498, 2016.
10. M. Ghiassi, B. A. Kijowski, R. E. Devor, and M. I. Dessouky, "An application of multiple criteria decision making principles for planning machining operations," *IIE Transactions*, vol. 16, no. 2, pp. 106–114, 1984.
11. K. Musselman and J. Talavage, "A Tradeoff Cut Approach to Multiple Objective Optimization," *Operations Research*, vol. 28, no. 6, pp. 1424–1435, 1980.
12. Y. T. Azene, *Work roll system optimisation using thermal analysis and genetic algorithm*. PhD thesis, School of Applied Sciences, Cranfield University, May 2011.