Journal of Environmental Management 336 (2023) 117689

Contents lists available at ScienceDirect

Journal of Environmental Management

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

Research article

Assisting decision-makers select multi-dimensionally efficient infrastructure designs – Application to urban drainage systems

Omid Seyedashraf[a,d], Andrea Bottacin-Busolin[a,c,*], Julien J. Harou[a,b]

[a] Department of Mechanical, Aerospace and Civil Engineering, The University of Manchester, Sackville Street, M13 9PL, Manchester, UK
[b] Department of Civil, Environmental & Geomatic Engineering, University College London, Gower Street, London, WC1E 6BT, UK
[c] Department of Industrial Engineering, University of Padova, Via Venezia 1, 35121, Padova, Italy
[d] Department of Civil Engineering, Kermanshah University of Technology, Kermanshah, Iran

ARTICLE INFO

Handling Editor: Jason Michael Evans

Keywords:
Multi-objective design
Pareto-front representation
Sustainable urban drainage systems
Fuzzy data clustering

ABSTRACT

Multi-objective design approaches can help identify future infrastructure system designs that appropriately balance different engineering, environmental, and other societal goals. Planners benefit from assessing the trade-offs implied by the best-performing infrastructure system solutions. However, a large number of possible efficient system designs, obtained when using multi-objective optimization, can be overwhelming to interpret. This study attempts to aid decision-making in multi-criteria infrastructure system design by reducing the complexity of the identified set of efficient infrastructure designs, i.e., the Pareto-front. A soft clustering algorithm is applied, which identifies similarities between solutions, partitions the front accordingly, and selects a set of representative solutions while preserving the multi-dimensional structure of the solutions on the efficiency frontier. Three post-optimization decision-making metrics are introduced to help quantify the overall performance of the Pareto-optimal designs to further summarize design process outputs for decision-makers. We apply the method to an illustrious urban drainage network case study. Results show how the approach can simplify Pareto-fronts with thousands of solutions into sets of highlighted designs that aid interpreting the trade-offs implied by the best-performing simulated systems.

1. Introduction

The design of urban infrastructure generally benefits from considering several goals. Whilst informative, this raises the issue of how to select a design given multiple performance criteria and the many possible different packages (combinations) of infrastructure options to consider. Heuristic multi-objective optimization algorithms (Maier et al., 2019) have been widely used by researchers for over a decade to tackle this task. This is due to their ability in dealing with multiple objectives to be minimized or maximized at the same time, resulting in a set of Pareto-optimal solutions ('Pareto-front'). In this context, a Pareto-front refers to a set of efficient solutions, in which there is not even a single solution with one better objective and other non-inferior objectives. This special decision-relevant set of designs contains all the future systems where performance cannot be improved in any dimension without simultaneously reducing performance in one or more other dimensions. The term 'many-objective optimization' is also sometimes used for design problems with more than three objectives (Fleming et al., 2005) where the heuristic search is used. This is *a posteriori* optimization, where planners evaluate the trade-offs between the optimized design objectives and select one or more solutions based on their engineering insight, past experience, and evaluation of costs and benefits; there is no need *a priori* of providing different priorities or weights between objectives.

However, documented use of multi-objective design methods by practitioners is more limited. Selecting or justifying an appropriate design, given the trade-offs they imply between several design objectives, remains a complex task (Blasco et al., 2008). Its complexity can increase with the number of optimized objectives, especially when the agreement must be achieved between several decision-makers who may have conflicting objectives. This process can be difficult if decision-makers lack relevant expertise or have time constraints. Moreover, in optimization problems, there is a general expectation that the number of optimal solutions will increase as the optimizer model approaches the final solution set with small steps. In other words, the higher the accuracy of a search algorithm, the larger the number of

* Corresponding author. Department of Mechanical, Aerospace and Civil Engineering, The University of Manchester, Sackville Street, M13 9PL, Manchester, UK.
E-mail address: andrea.bottacinbusolin@unipd.it (A. Bottacin-Busolin).

https://doi.org/10.1016/j.jenvman.2023.117689
Received 14 October 2022; Received in revised form 15 January 2023; Accepted 6 March 2023
Available online 14 March 2023

solutions in the resultant Pareto-approximate set (Hadka and Reed, 2013), implying that a high degree of accuracy in an optimization problem corresponds to a denser sampling of the non-dominated solution space. Pareto-fronts with many solutions make it harder to visualize and interpret trade-offs between design objectives, making it more difficult for decision-makers to use results to inform their selection of a particular portfolio of options (a design). The psychology literature has pointed out that decision-makers have difficulties in processing large data sets and using them to take decisions (Duro et al., 2014; Kaplan, 1995).

Urban drainage infrastructure design is an example where this topic is of relevance. Urban drainage systems protect cities from the risk of flooding and protect their downstream environments and ecosystems from pollution (Butler and Davies, 1999). However, climate change and rapid urban expansion have made some existing drainage infrastructure insufficient to protect cities from intense rainfall events (Huang et al., 2020). Potential solutions include expansion and refurbishment of traditional grey drainage infrastructure (Barreto Cordero, 2012), real-time flow control of drainage systems (Abou Rjeily et al., 2018), and applying sustainable urban drainage systems (SuDS) (Elliott and Trowsdale, 2007; Li et al., 2022; Shojaeizadeh et al., 2021; Torres et al., 2020) to reduce surface runoff. The latter have received attention in recent years because of their multi-functional properties such as improvement of water quality in receiving water bodies by promoting sediment settling, filtering, and biological breakdown of pollutants (Woods Ballard et al., 2015), improving biodiversity (Wright, 2011), delivering recreational opportunities, improving the mental and physical health of the residents (Mell, 2010), and turning floods into alternative water supplies. The various advantages of using SuDS (co-benefits), compared to the first two methods, can complicate their design-related decision--making process.

Several studies have used multi-objective optimization methods for SuDS design (Eckart et al., 2018; Huang et al., 2022; Koc et al., 2021; Macro et al., 2019; Mani et al., 2019; Seyedashraf et al., 2021a). For example, Eckart et al. (2018) applied the Borg multi-objective evolutionary algorithm (MOEA) (Hadka and Reed, 2013) to a SuDS design problem in Windsor, Canada. The model was used to find efficient surface areas of four types of sustainable urban drainage facilities, including infiltration trenches, rain gardens, permeable pavements, and rain barrels. Seyedashraf et al. (2021a) found Pareto-optimal designs of SuDS to reduce risk of flooding in urban areas with different surface slopes. The promotion of adaptivity and resilience of urban drainage infrastructure services, in response to variabilities in future climate, urbanization, and population growth, has also resulted in several studies in this field. For instance, Casal-Campos et al. (2018) considered model robustness in delivering long-term and resilient sustainable urban drainage infrastructure services. Babovic and Mijic (2019) used an adaptation tipping points approach to investigate how urban drainage systems respond to climate variability, in terms of depth and intensity of future rainfall events, in the Cranbrook urban catchment, London.

Although there are several studies on the use of multi-objective optimization models to assist the design of urban drainage systems, the literature on aiding the use and interpretation of optimized design results in the decision-making process is sparse. In the area of designing urban water networks, this paper's intended contribution is to help narrow the number of optimization outputs down to a handful of solutions while preserving the distributional structure of the full Pareto-fronts, thereby making them easier to interpret and use. An illustrative urban drainage system was considered for generating a sample of efficient system designs, and a soft clustering algorithm was used, which partitions the Pareto-front into a number of clusters, each of which includes solutions that behave similarly in terms of fulfilling the design goals. A representative solution is then assigned to each cluster, making it possible to reduce a Pareto-front with several solutions to a smaller number of similar infrastructure designs. In this way, the representative solutions highlight designs that behave roughly similarly, hence preserving the distributional structure of the Pareto-front like its original version, yet helping stakeholders identify and differentiate between subgroups of similar efficient designs. Moreover, since overall performance metrics can in some cases assist decision-makers (Duro et al., 2014), we introduced three post-optimization metrics that look at classes of data sets to evaluate the overall performance of the efficient designs. These can be used when decision-makers have achieved their major design goals and need an objective way to distinguish between a shortlist of designs. The metrics were used to rank clusterized solutions as a decision-making aid when there is no further preference for any of the design objectives because the essential design requirements have been met.

2. Methodology

The proposed framework seeks both to simplify the analysis of multi-objective optimization outputs and help to bring consistency in decision-making when planners might be overwhelmed by the number of different combinations of design options. It aims to simplify decision-making in multi-objective urban water infrastructure design problems by reducing the complexity of Pareto-fronts and helping decision-makers focus their attention on groups of similar system designs rather than having to evaluate an overwhelming number of marginally different designs.

A simulation-optimization framework was applied with a set of decision-making tools to evaluate the resultant Pareto-front. Flow routing simulations were carried out using the Storm Water Management Model (SWMM) (Gironás et al., 2010), whereas Borg MOEA was employed for model optimization and to generate sets of SuDS designs with best trade-offs between the objective functions. The metaheuristic multi-objective search process we use is *a posteriori* optimization, where prior preferences on design objectives are not defined, and decisions are made based on the performance trade-offs implied by efficient solutions of the Pareto-set (Coello et al., 2007). A fuzzy data clustering algorithm was used to explore data structure in the resultant Pareto-front and classify designs in a trial-and-error fashion by calculating the global silhouette index for each clustering scheme as a measure to evaluate quality of the clusters. Subsequently, additional metrics were used to quantify the overall performance of the representative solutions.

2.1. Many-objective optimization

In many-objective optimization problems, one or several search algorithm/s may be implemented to reduce the decision space to a set of solutions that maximize and/or minimize multiple design objectives subject to a given set of design constraints (Maier et al., 2019). This process can be mathematically described as follows:

$$\underset{\boldsymbol{l}\ \varepsilon\ \boldsymbol{\varphi}}{\text{Minimize}} : \boldsymbol{F}(\boldsymbol{l}) = (F_1(\boldsymbol{l}), F_2(\boldsymbol{l}), \ldots, F_{N_o}(\boldsymbol{l})) \tag{1}$$

$$\text{Subject to} : \begin{cases} C_{\text{eq},j}(\boldsymbol{l}) = 0 j = 1, \ldots, N_q \\ C_{\text{in},k}(\boldsymbol{l}) \leq 0 k = 1, \ldots, N_r \end{cases} \tag{2}$$

where $\boldsymbol{F}(\boldsymbol{l})$ is a vector of objective functions, $F_i(\boldsymbol{l})$, which characterizes performance of the vector of decision variables, $\boldsymbol{l}$, with N_o number of objectives in the decision space, $\boldsymbol{\varphi}$. Moreover, $C_{\text{eq},j}(.)$ and $C_{\text{in},k}(.)$ are equality and inequality functions with N_q and N_r constraints, respectively.

2.2. Data clustering

Data clustering algorithms help discover groupings of data points in large datasets based on measures of similarity. They can be used to organize, compress, and categorize large amounts of data for diverse applications, including image segmentation, text mining, speech recognition, and health monitoring (Bezdek et al., 1984; Mitra et al., 2004; Ng et al., 2006; Satour et al., 2020). In the past years, such algorithms have also been applied in the context of urban water infrastructure design to investigate spatial characteristics of system components (Huang et al., 2015; Liu et al., 2016; Muhammed et al., 2017). Various clustering algorithms have been introduced based on the three main data partitioning approaches, including: (1) hierarchical, e.g. clustering using representatives (CURE) (Guha et al., 1998); (2) exclusive, e.g. k-means clustering algorithm (MacQueen, 1967); and (3) overlapping clustering approaches, such as the fuzzy c-means clustering algorithm (Bezdek, 1981). The hierarchical clustering algorithms operate by successive clustering of initially partitioned data points via a dendrogram in which the root of all subsets corresponds to the main data set. The exclusive clustering algorithms partition datasets into a number of groups with crisp boundaries where each data point belongs to only one cluster (MacQueen, 1967). The overlapping clustering algorithms, however, allocate membership grades to each data point allowing them to fit in multiple groups (Bezdek, 1981; Bezdek et al., 1984; Zadeh, 1965), for which the membership values can range between 0 and 1 and sum up to 1 for all clusters. The closer the membership values to 1 the better the clustering process. In this category, the fuzzy c-means clustering algorithm is the most widely used method, which groups similar data objects into clusters.

In the fuzzy c-mean data clustering method, the final groupings of a dataset into different groups are not as clear as in the case of k-means clustering where each data in the dataset is assigned to only one cluster. In fuzzy c-means clustering, however, each object belongs to all clusters to a degree of membership. Fuzzy c-means is based on the fuzzy theory where the membership value of an object is not explicitly assigned to a value of 1 or 0, representing a member or not of a cluster, respectively. Using this technique, the centroid of each cluster corresponds to the center of the data to which each object contributes with its own degree of membership. The algorithm operates by selecting and iteratively updating hypothetical cluster centers, C_k as follows:

$$C_k = \frac{\sum_{i=1}^{N_c} (dm_{ik})^w \times D_i}{\sum_{i=1}^{N_c} (dm_{ik})^w} \tag{3}$$

so as to minimize a loss function, L,

$$L = \sum_{i=1}^{N_c} \sum_{k=1}^{N_p} (dm_{ik})^w \times R_{ik} \tag{4}$$

subjected to the following constraint (Bezdek, 1981):

$$\sum_{k=1}^{N_s} dm_{ik} = 1 \tag{5}$$

where dm_{ik} is the membership degree of the i^{th} data point to the k^{th} cluster, N_p is the number of data points, N_c is the number of clusters, w is a weighting factor, D_i is the i^{th} data point, and R_{ik} is the distance between the i^{th} data point and k^{th} cluster center, which must be minimized by an integrated optimization model.

Here, the degree of membership depends on the closeness of each data object to the respective cluster center, hence, the higher the degree of membership value of a data object the greater its association with a particular cluster center. The fuzzy c-means clustering algorithm was used in this study with its weighting factor and number of clusters to be determined based on a trial-and-error method that calculates the global silhouette index for each clustering scheme as a measure to evaluate quality of the generated clusters (Sun et al., 2015; Zio and Bazzo, 2011). The silhouette index, s, ranges between -1 and $+1$, where $s = +1$ implies that a solution is distant from other designs in the nearest cluster, $s = -1$ indicates that the solution is assigned to a wrong partition, and $s = 0$ implies that the solution is not distinctly assigned to a particular partition. Accordingly, the larger the silhouette value, the better the clustering and consistency among the cluster data objects.

The global silhouette index, GS, is calculated as follows:

$$GS = \frac{1}{N_c} \sum_{j=1}^{N_c} \left(\frac{1}{N_e} \sum_{i=1}^{N_e} s(i) \right) \tag{6}$$

where N_e is the number of SuDS designs in each cluster, and $s(i)$ is the silhouette index of the i^{th} design solution defined as follows:

$$s(i) = \frac{D_a(i) - D_b(i)}{\max\{D_a(i), D_b(i)\}} \tag{7}$$

where $D_a(i)$ and $D_b(i)$ are average distances of the i^{th} solution from other solutions in the same and nearest clusters, respectively.

2.3. Ranking design portfolios

While the decision-making process is often based on subjective views and previous experiences (Duro et al., 2014), the assessment of the performance of the optimized designs can be aided by decision support tools. In this study, three decision support metrics were used to provide a measure of the overall performance of the Pareto-optimal solutions. Once planners have achieved their major objectives but several possible designs are still left, an objective way to further distinguish between a shortlist of designs can be helpful. Generally, the measures can be defined based on normalized distances of objective function values from the minimum values obtained. Accordingly, the overall performance of each design can be quantified using the 1-norm ($\|N\|_1$), 2-norm ($\|N\|_2$), and the infinity norm ($\|N\|_\infty$) of the vectors of normalized objectives (Reynoso-Meza et al., 2013; Sánchez-Orgaz et al., 2015).

In this regard, the 1-norm $\|N\|_1$ corresponds to the summation of the normalized objectives for a particular design solution (Zio and Bazzo, 2011), and is calculated as follows:

$$\|N\|_1 = \sum_{k=1}^{N_o} |\overline{F}_k| \tag{8}$$

where N_o is the number of designs in the reduced front and the k^{th} normalized objective in the i^{th} solution, $\overline{F}_k(i)$, is defined as:

$$\overline{F}_k(i) = \frac{F_k(i) - F_{k,\min}}{F_{k,\max} - F_{k,\min}}, k = 1, 2, \ldots, 4 and i = 1, 2, \ldots, 5 \tag{9}$$

where $F_k(i)$ is the k^{th} objective of the i^{th} SuDS solution and $F_{k,min}$ and $F_{k,max}$ are the maximum and minimum values of the k^{th} objective in the Pareto-front.

Alternatively, the 2-norm can be used to find the normalized Euclidian distance between a point in the objective function space and the point corresponding to the best objective function values obtained. The 2-norm is calculated as follows:

$$\|N\|_2 = \sqrt{\sum_{i=k}^{M} |\overline{F}_k|^2} \tag{10}$$

Another performance metric can be defined using the infinite norm, which gives the maximum value of the normalized objective functions in the available designs. This metric is helpful to find a solution with the least-worst objective function values, i.e.:

$$\|N\|_\infty = \max \left(|\overline{F}_1|, |\overline{F}_2|, |\overline{F}_3|, |\overline{F}_4|\right) \tag{11}$$

When put side by side in a parallel axes plot, these metrics can help decision-makers evaluate the overall performance of each solution and identify the best solutions within their design preferences

Fig. 1 summarizes the proposed workflow for many-objective optimization of SuDS. This involves using SWMM for evaluating system

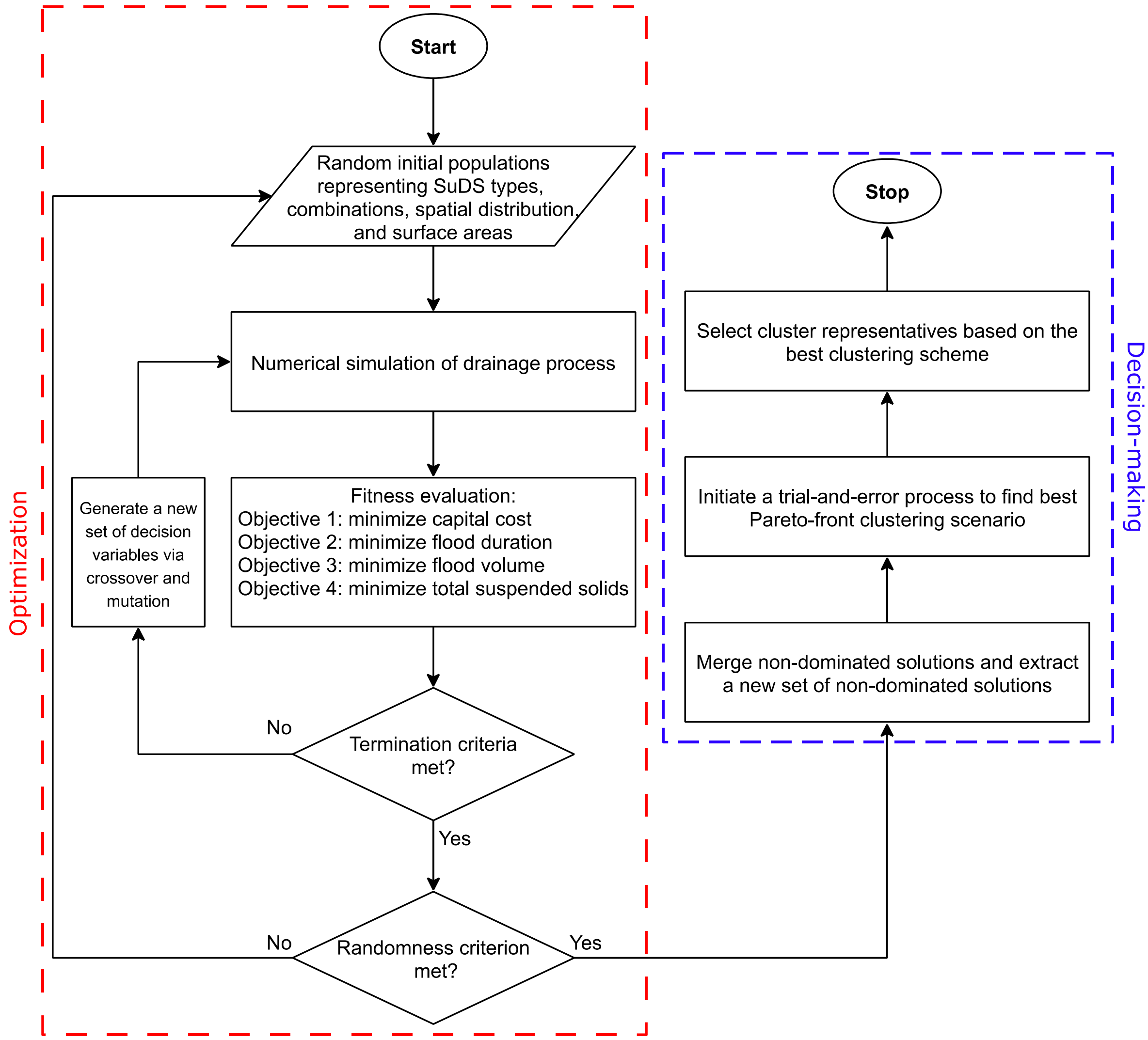

Fig. 1. Flowchart of the proposed many-objective optimization approach for design of sustainable urban drainage infrastructure.

performance coupled with the Borg MOEA to obtain a set of Pareto-optimal SuDS designs, and with the fuzzy c-mean clustering algorithm to determine a smaller set of representative solutions. In the latter stage, additional metrics were used to quantify the overall performance of the representative solutions.

3. Application of the proposed approach

3.1. Case study

An illustrative urban drainage system, comprising 7 subcatchments, 11 junctions, and 11 conduits (Fig. 2), was considered to demonstrate application of the proposed decision-making framework.

This case study is a modified version of the model presented by Lewis A Rossman (2017) and used as a standard benchmark in different drainage design studies considering urban flooding both in terms of water quantity and quality (Giacomoni and Joseph, 2017; Nehrke and Roesner, 2002, 2004; Sambito et al., 2020). To reduce model complexity, it was assumed that there are no restrictions on placing different types of sustainable urban drainage components in the sub-catchments. Moreover, the pipe diameters were halved, compared to the original case study presented by Rossman (2017), in order to generate a scenario where additional drainage infrastructure would be needed to avoid flooding in the region of interest.

SWMM (Gironás et al., 2010) was used to simulate drainage processes in the study area. SWMM incorporates three flow routing models, including steady flow, kinematic wave, and dynamic wave models, to simulate the flow of runoff through the drainage network (Rossman, 2017). The dynamic wave flow routing model was used in this study due to its ability to reproduce pressurized and backwater flow conditions by solving the Saint-Venant equations (Meza and Oliva, 2003). SWMM can also simulate pollution build-up and transport; and in its latest versions, it allows simulation of best management practices such as rain gardens, bio-retention cells, green roofs, permeable pavements, infiltration trenches, rain barrels, rooftop disconnections, and vegetative swales (Gironás et al., 2010). In this study, the first six sustainable urban drainage components were used to decrease stormwater runoff, however, further preferences to certain types of SuDS may be applied by the decision-makers in terms of optimization constraints throughout the search process.

3.2. Optimization model

A many-objective optimization and decision-making approach developed to find efficient drainage system designs where the decision variables include types, combinations, surface areas, and spatial distribution of SuDS components in the catchment. The search algorithm considers simultaneous minimization of average flood duration, total

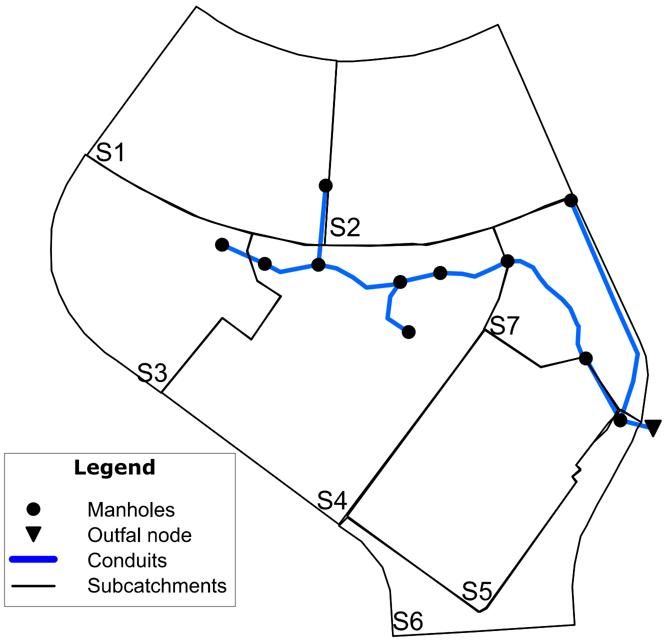

Fig. 2. Schematic map of the synthetic case study.

flood volume, total suspended solids (TSS), and capital cost of SuDS components for balanced design objectives. Accordingly, the optimization objectives were considered as follows:

$$\text{Minimize } \boldsymbol{F}(\boldsymbol{l}) = (F_C(\boldsymbol{l}), F_{FD}(\boldsymbol{l}), F_{FV}(\boldsymbol{l}), F_{TSS}(\boldsymbol{l})) \tag{12}$$

where F_C is capital cost, F_{FD} is average flood duration, F_{FV} is overall flood volume, and F_{TSS} is TSS load at system outfall. The capital cost function is defined as:

$$F_C = \sum_{i=1}^{N_s} \sum_{j=1}^{2} (c_{ij} \times a_{ij}) \tag{13}$$

where N_s is the number of subcatchments, a_{ij} and c_{ij} are respectively the surface area and capital cost of each SuDS component extracted from cost databases published by the Washington State Department of Ecology & Herrera Environmental Consultants (2012) and online vendors.

The overall TSS at the system outfall was extracted from numerical simulations whereas the average flood duration, F_{FD}, and overall flood volume, F_{FV}, for each SuDS design were calculated as follows:

$$F_{FD} = \frac{\sum_{i=1}^{N_j} fd_i}{n_j} \tag{14}$$

$$F_{FV} = \sum_{i=1}^{N_j} fv_i \tag{15}$$

where N_j is the number of system junctions, and fd_i and fv_i are flood duration and flood volume for each system junction, respectively.

In this study, the Borg MOEA was used (Hadka and Reed, 2013) for the optimization due to its proven performance in dealing with multi-criteria water system design, planning, and management (Li et al., 2021; Seyedashraf et al., 2021b; Zatarain Salazar et al., 2017). Borg benefits from a combination of various search algorithms, which operate by evolving an initial population of solutions towards solutions with higher fitness values. Moreover, a random seed analysis was performed by a total of 30 optimization runs each of which started with a random initial population to ensure a sufficient level of diversity in the final set of Pareto-optimal solutions. To verify convergence of the optimization process, the evolution of the hypervolume indicator was evaluated for each optimization run against the number of objective function evaluations across the initial populations (Appendix A).

4. Results and discussion

This section illustrates the results obtained from applying the proposed multi-criteria system design assistance method to an illustrative case study involving sustainable urban drainage infrastructure. A real-world case study might have better helped demonstrate its ability to assist decision-making but this case study system, widely used in the literature as a benchmark provides an accessible example of use (Nehrke and Roesner, 2002; Sambito et al., 2020; Seyedashraf et al., 2022).

4.1. Pareto-optimal solutions

A reference set was generated by combining solutions from optimization runs resulting in the roughly 15,000 Pareto-optimal solutions shown in a parallel coordinate plot (Fig. 3). Parallel axis plots (Inselberg, 2009) (also named parallel coordinate plots) have been widely used in multi-objective optimization problems relating to water management and water infrastructure design to support decision-making and exploration of relationships between design goals (Matteo et al., 2016; Seyedashraf et al., 2021a). Here, the vertical axes represent design objectives and each colored line connecting the axes represents an infrastructure system design implying different trade-offs between the design objectives.

4.2. Clustering the pareto-front

In this study, a decision-making framework for many-objective optimization of SuDS is proposed that narrows down the Pareto (efficient) fronts to a handful of designs while preserving their distributional structure. The data clustering technique applied here groups similar SuDS design according to their performances and in terms of the previously defined objective functions. Here, we assume that the designs are acceptable to planners if the capital cost is less than 1.5 M$. Accordingly, any arrangement of weighting factors and numbers of clusters covering this range may be considered suitable to narrow down the Pareto-front. The number of data clusters can be determined by experience or based on the expertise/knowledge of the decision-makers or even determined automatically by an optimization method or trial-and-error. In this study, it is assumed that decision-makers have agreed to analyze 5 SuDS designs in the Pareto-front. Fig. 4 illustrates the trial-and-error process carried out to select the best fuzzy c-means clustering scheme with a maximum possible global silhouette index, which in this case was 0.68 for 5 data clusters and a weighting factor of 2. In this figure, marker sizes represent the global silhouette value for each trial, where the larger a marker the more successful the clustering scheme.

Fig. 5 depicts the 2D scatter plots of the primary Pareto-front (shown in Fig. 3) along with their cluster centers (hypothetical points) and cluster representatives (the existing designs nearest to cluster centers). Here, the cluster centers are synthetic data objects that do not belong to the Pareto-front, yet, imply where a good representative of a cluster can be. The designs closest to the cluster centers (red circle markers in Fig. 5) were selected as cluster representatives (black triangle markers in Fig. 5). The figure shows the degree of consistency among data points in each cluster where the larger the silhouette value, the better the clustering. In this figure, the color range indicates the extent to which the designs belong to each cluster.

It can be seen in Fig. 5 that the fuzzy c-means algorithm used in this study has successfully clustered solutions according to their performance relative to each objective function. This is especially noticeable in the scatter plots in Fig. 5 illustrating the trade-offs between capital costs and flood volumes, capital costs and TSS values, flood volumes and TSS values, and flood volumes and average flood duration. However,

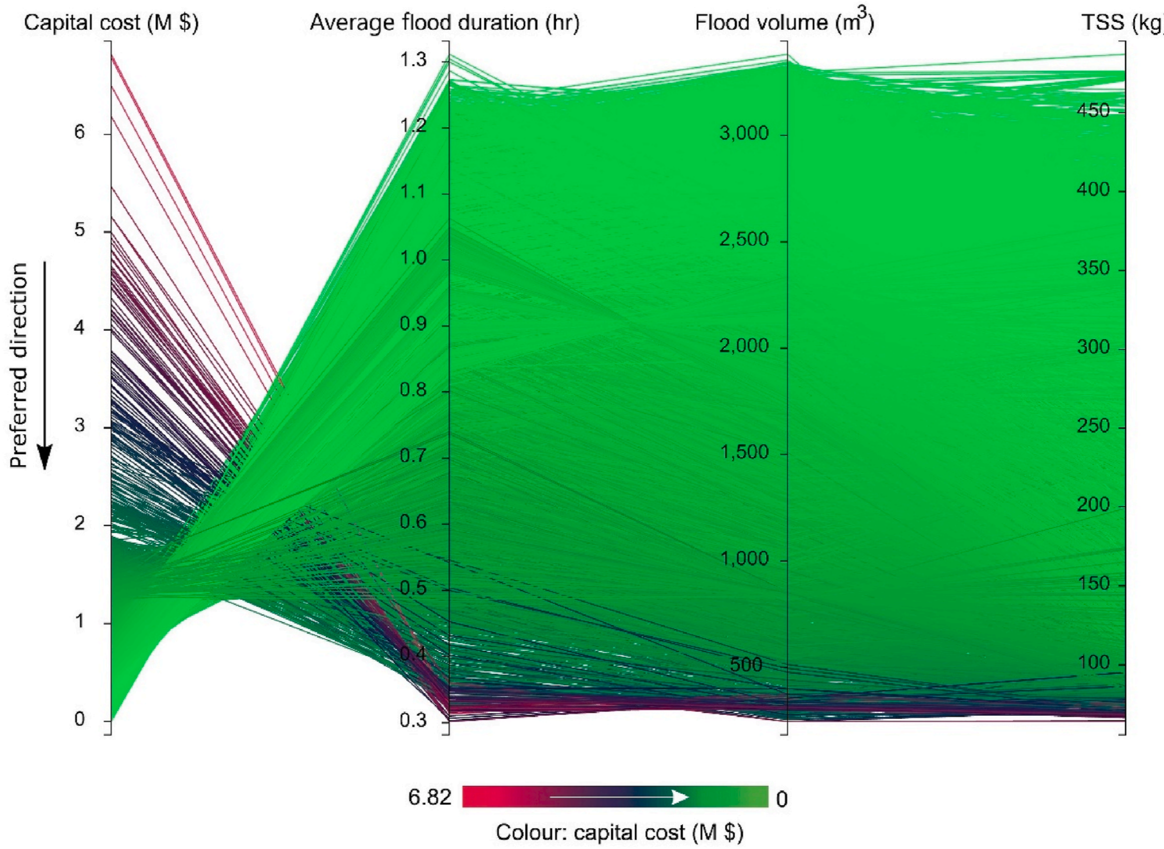

Fig. 3. Parallel axis plot of around 15,000 Pareto-optimal sustainable urban drainage system designs showing trade-offs between the optimization objectives, including capital cost, flood volume, average flood duration, and TSS. Here, the arrows show the direction of preference, the vertical axes represent design objectives, and each colored line connecting the axes represents a sustainable urban drainage infrastructure design implying different trade-offs between the design objectives.

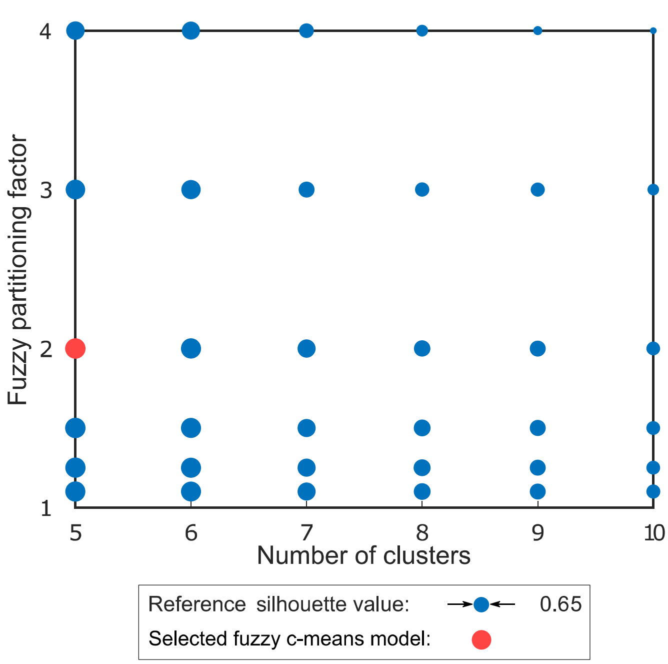

Fig. 4. The trial-and-error process used to find an efficient Pareto-front clustering scheme to compress and categorize Pareto-optimal designs of sustainable urban drainage infrastructure. Here, marker sizes represent their global silhouette values where the larger a marker the more successful the clustering scheme.

shortcomings can be seen in clusters representing trade-offs between the average flood durations and TSS values, as well as capital costs and average flood durations, which can be inevitable when dealing with many-objective optimization problems with conflicting design objectives.

Fig. 6 shows the trade-offs between optimized performance objectives of the cluster representatives. In this figure, selection filters are defined as grey boxes which act as post-optimization constraints to reflect decision-makers' preferences and/or requirements by removing undesired solutions from further evaluations (the www.polyvis.org website allows making interactive parallel coordinate plots). Here, each colored line represents a particular cluster representative solution, dashed lines are SuDS designs that do not fit in the post-optimization constraints, while the solid line (yellow in this case) represents a design that meets the filters imposed by decision-makers.

A linear relationship between flood volume and TSS as well as a conflicting trade-off between capital costs and other three objective functions are evident in Fig. 6. Using the proposed Pareto-front simplification approach allows decision-makers easy exploration of the solutions as well as the trade-offs between optimization objectives, thus providing decision-makers easy-to-understand information about the solution space. Moreover, from Figs. 3 and 6 it follows that when working on a subset of multivariate data in the design of urban water systems, Pareto-front clustering in a parallel coordinate plot can help decision-makers better focus their exploration of efficient alternative designs. The decision-making process can be further aided by decision support that ranks the overall performance of each design once planners have already achieved their major design goals. Here the 1-norm ($\|N\|_1$), 2-norm ($\|N\|_2$), and the infinity norm ($\|N\|_\infty$) of the vectors of normalized objectives were applied to help decision-makers evaluate the

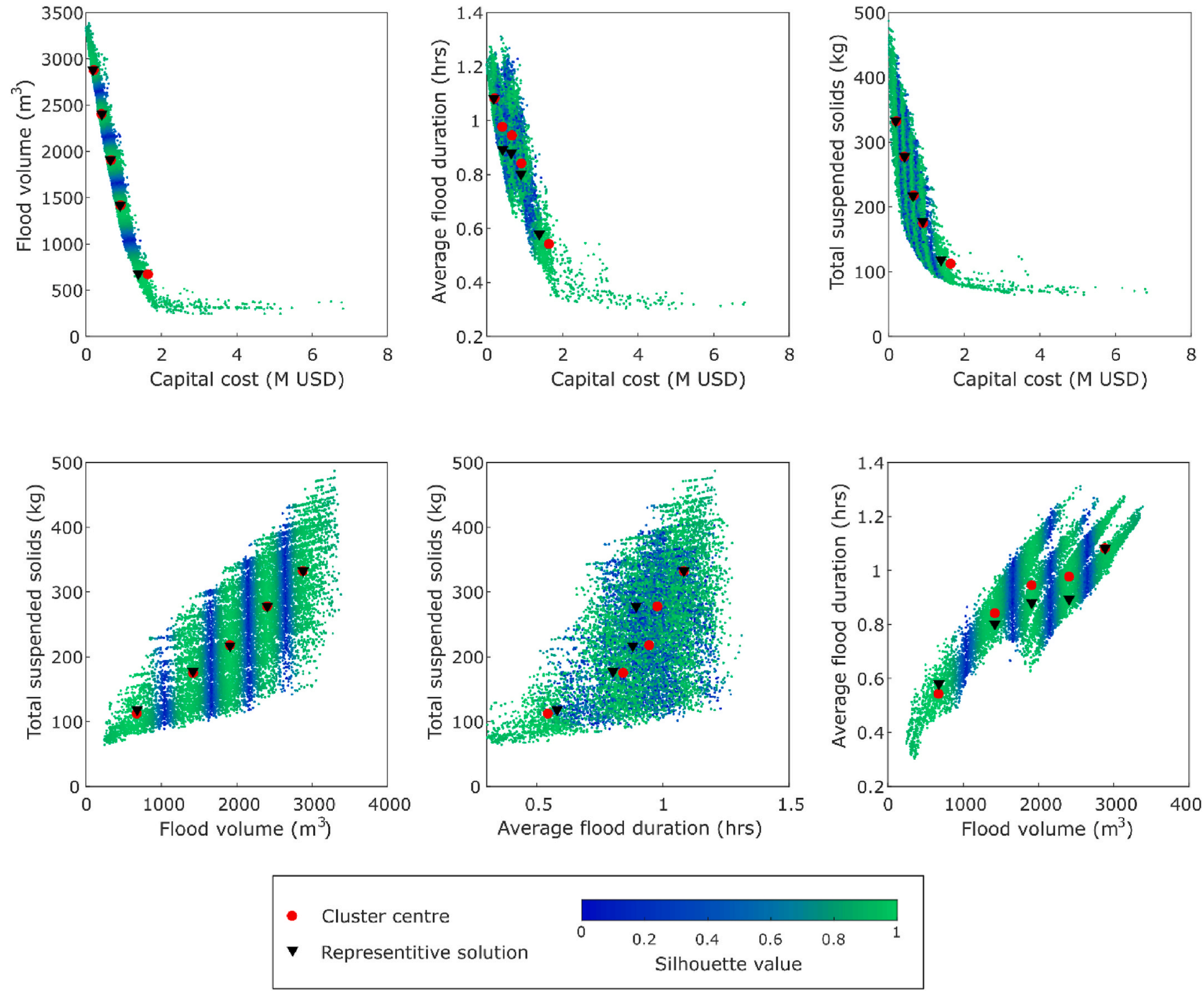

Fig. 5. 2D scatter plot of Pareto-optimal designs of sustainable urban drainage infrastructure along with their cluster centers (red circle markers) and cluster representatives (black triangle markers). The color range represents silhouette indices of the solutions according to the clusters they belong to. The figure shows the degree of consistency among data points in each cluster where the larger the silhouette value, the better the clustering.

overall performance of each solution and rank solutions that all fall within their preferred design parameter preferences (Fig. 7).

Fig. 7 can help decision-makers select SuDS design alternatives that fulfill the project requirements while providing three measures of the overall performance of the Pareto-optimal solutions. According to the results, the 3rd cluster representative is a relatively economic design, and its non-monetary objectives roughly fall in the middle of the vertical axes, which is consistent with its 1-norm value. Although the 3rd and the 5th cluster representatives almost share identical 2-norm values, there is a noticeable difference between their infinity norm values as the 5th solution has at least one worst-case objective value. Moreover, the 2nd solution comprises objective values with the second-highest values of flood volume and infinity norm. Furthermore, the 4th solution is 40% more costly compared to the 3rd solution and performs best with respect to non-monetary objectives while sharing an identical infinity norm of 0.60.

5. Conclusions

Decision-makers may find it difficult to process and make decisions based on the results of design studies that produce large numbers of alternative designs. The new crop of multi-objective *a posteriori* optimization methods used in infrastructure system engineering fit in this category by typically resulting in large numbers of alternative 'efficient' designs. In this paper, we propose a new framework for narrowing down the number of Pareto-optimal infrastructure designs to a more parsimonious summarized set while preserving the multi-dimensional structure of the full set of optimized solutions. To this end, a soft clustering algorithm was used to aggregate Pareto-optimal solutions based on performance similarities. The algorithm was able to narrow down a Pareto-front of around 15,000 efficient sustainable urban drainage system designs to 5 designs within the range of acceptable cost and benefit. For each cluster, the design closest to each virtual cluster center is chosen as the representative solution of the cluster. The proposed method can help urban drainage infrastructure planners more easily and effectively interpret and learn from efficient system designs and the trade-offs and synergies they imply. The evaluation of the performance of different design solutions can be further complicated by the need to consider the effects of parameter uncertainty and climate change, which were not addressed in this study.

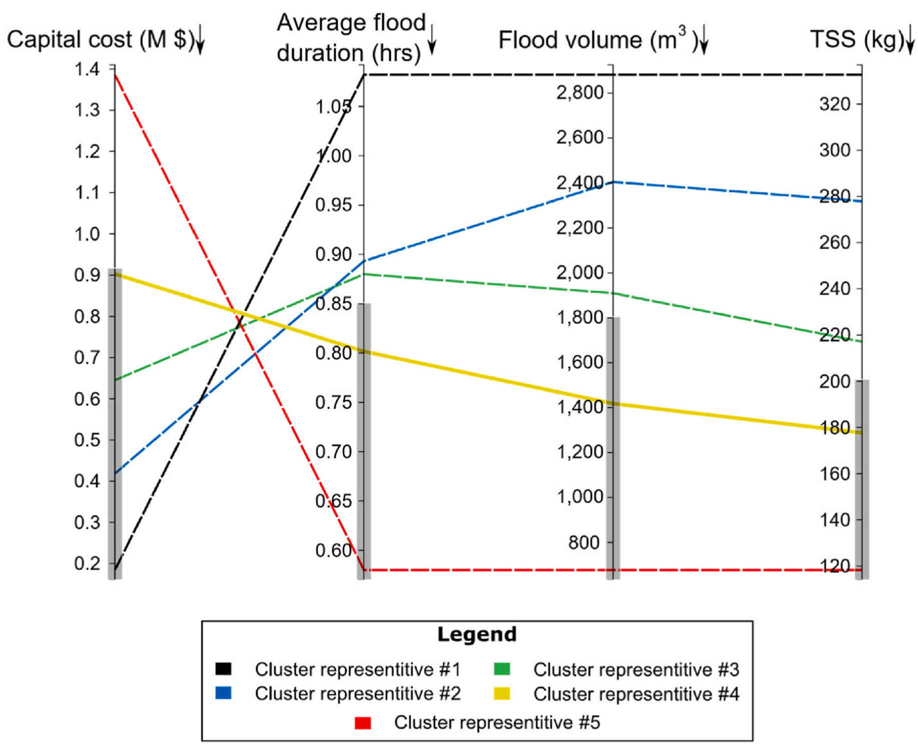

Fig. 6. Parallel axes plot of the reduced Pareto-front where each colored line represents a particular cluster representative, (existing designs nearest to the cluster centers) dashed lines are sustainable urban drainage system designs that do not fit in the post-optimization constraints (the grey 'filter' boxes that stakeholders can impose on the axes interactively) while the solid line is the preferred solution in this case. The arrows show the direction of preference and grey boxes represent filters to eliminate undesirable solutions from further analysis.

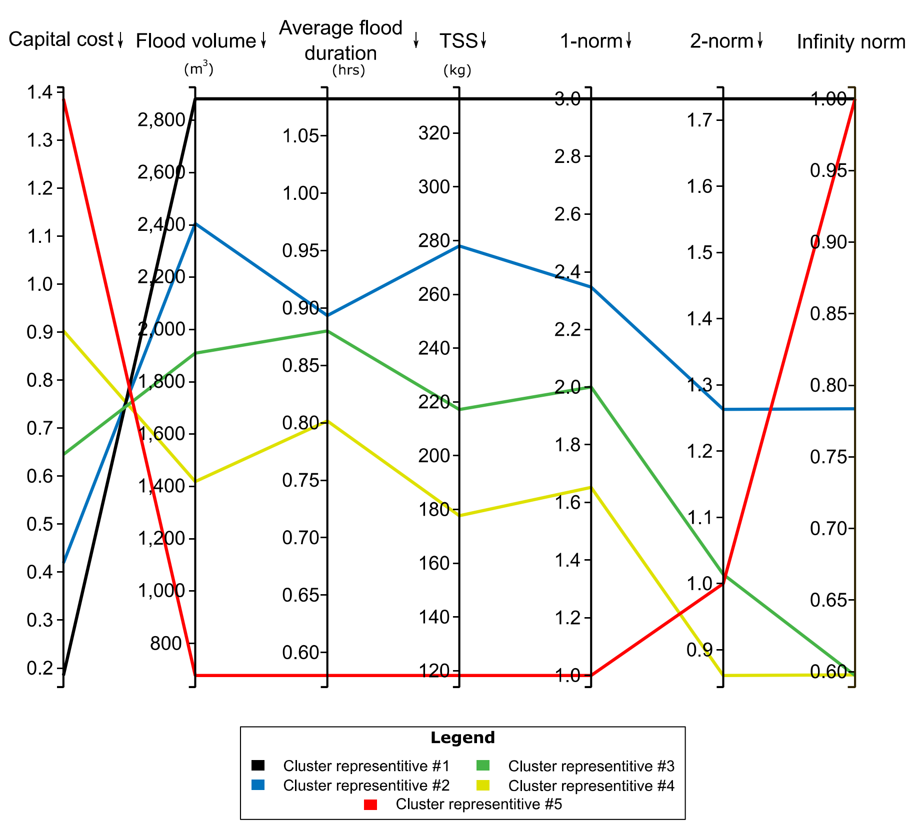

Fig. 7. Parallel axes visualization of Pareto-optimal sustainable urban drainage infrastructure designs. This figure illustrates trade-offs between the design objectives and ranks each solution in terms of its 1-norm, 2-norm, and infinity norm values. The arrows show the direction of preference, and each colored line represents a sustainable urban drainage system design.

Credit author statement

Omid Seyedashraf: Conceptualization, Methodology, Software, Visualization, Writing – Original draft preparation, Writing – Review and Editing. Andrea Bottacin-Busolin: Supervision, Writing – Review and Editing, Funding acquisition. Julien Harou: Supervision, Writing –

Review and Editing, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Omid Seyedashraf reports financial support was provided by Thames Water Utilities Ltd. Omid Seyedashraf reports a relationship with Thames Water Utilities Limited that includes: funding grants.

Data availability

The data that has been used is confidential.

Acknowledgments

This work was jointly supported by Thames Water Utilities Ltd and the University of Manchester.

Appendix A. Convergence of the optimization process

In this study, 30 many-objective sustainable urban drainage infrastructure optimizations were carried out each of which initialized with different random populations to reduce randomness dependency of the implemented search algorithms. Fig. A1 depicts the hypervolume evolution across initial populations over the number of objective function evaluations. The shaded area bounds the hypervolume indicators of the Pareto-fronts and the solid curve represents the mean hypervolume of the different runs. The hypervolume indicator is a commonly used quality measure for evaluating the performance of multi-objective search algorithms. According to the hypervolume evolution in this figure, the Pareto-front stabilizes after completing around 20,000 objective function evaluations.

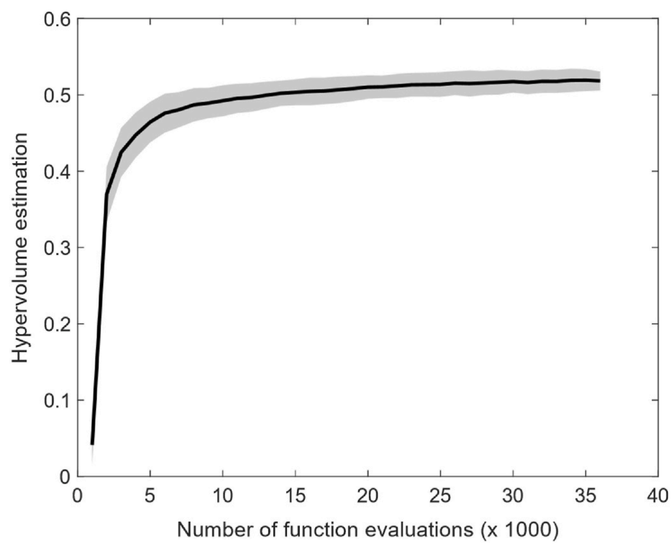

Fig. A1. Hypervolume evolution across 30 random initial populations over the number of objective function evaluations. The shaded area bounds the hypervolume of the Pareto-optimal sustainable urban drainage infrastructure designs and the solid line represents the mean hypervolume indicator reached through the experiments.

References

Abou Rjeily, Y., Abbas, O., Sadek, M., Shahrour, I., Hage Chehade, F., 2018. Model predictive control for optimising the operation of urban drainage systems. J. Hydrol. 566, 558–565. https://doi.org/10.1016/j.jhydrol.2018.09.044.

Babovic, F., Mijic, A., 2019. The development of adaptation pathways for the long-term planning of urban drainage systems. J. Flood Risk Manag. 12, e12538.

Barreto Cordero, W.J., 2012. Multi-objective Optimization for Urban Drainage Rehabilitation. CRC/Balkema.

Bezdek, J.C., 1981. Pattern recognition with fuzzy objective function algorithms. In: Pattern Recognition with Fuzzy Objective Function Algorithms. https://doi.org/10.1007/978-1-4757-0450-1.

Bezdek, J.C., Ehrlich, R., Full, W., 1984. FCM: the fuzzy c-means clustering algorithm. Comput. Geosci. 10, 191–203. https://doi.org/10.1016/0098-3004(84)90020-7.

Blasco, X., Herrero, J.M., Sanchis, J., Martínez, M., 2008. A new graphical visualization of n-dimensional Pareto front for decision-making in multiobjective optimization. Inf. Sci. 178, 3908–3924. https://doi.org/10.1016/j.ins.2008.06.010.

Butler, D., Davies, J., 1999. Urban Drainage, Urban Drainage. CRC Press, London. https://doi.org/10.1201/b13173.

Casal-Campos, A., Sadr, S.M.K., Fu, G., Butler, D., 2018. Reliable, resilient and sustainable urban drainage systems: an analysis of robustness under deep uncertainty. Environ. Sci. Technol. 52, 9008–9021. https://doi.org/10.1021/acs.est.8b01193.

Coello, C.A.C., Lamont, G.B., Van Veldhuizen, D.A., Coello Coello, C.A., Lamont, G.B., Van Veldhuizen, D.A., 2007. Evolutionary Algorithms for Solving Multi-Objective Problems, Evolutionary Algorithms for Solving Multi-Objective Problems. Springer, New York, NY. https://doi.org/10.1007/s11276-014-0817-8.

Duro, J.A., Kumar Saxena, D., Deb, K., Zhang, Q., 2014. Machine learning based decision support for many-objective optimization problems. Neurocomputing 146, 30–47. https://doi.org/10.1016/j.neucom.2014.06.076.

Eckart, K., McPhee, Z., Bolisetti, T., 2018. Multiobjective optimization of low impact development stormwater controls. J. Hydrol. Elsevier B.V. https://doi.org/10.1016/j.jhydrol.2018.04.068.

Elliott, A.H., Trowsdale, S.A., 2007. A review of models for low impact urban stormwater drainage. Environ. Model. Software 22, 394–405. https://doi.org/10.1016/j.envsoft.2005.12.005.

Fleming, P.J., Purshouse, R.C., Lygoe, R.J., 2005. Many-objective optimization: an engineering design perspective. In: Lecture Notes in Computer Science. https://doi.org/10.1007/978-3-540-31880-4_2.

Giacomoni, M.H., Joseph, J., 2017. Multi-objective evolutionary optimization and Monte Carlo simulation for placement of low impact development in the catchment scale. J. Water Resour. Plann. Manag. 143, 04017053 https://doi.org/10.1061/(ASCE)WR.1943-5452.0000812.

Gironás, J., Roesner, L.A., Rossman, L.A., Davis, J., 2010. A new applications manual for the storm water management model (SWMM). Environ. Model. Software 25, 813–814. https://doi.org/10.1016/j.envsoft.2009.11.009.

Guha, S., Rastogi, R., Shim, K., 1998. CURE: an efficient clustering algorithm for large databases. SIGMOD Rec. 27, 73–84. https://doi.org/10.1145/276305.276312.

Hadka, D., Reed, P., 2013. Borg: an auto-adaptive many-objective evolutionary computing framework. Evol. Comput. 21, 231–259. https://doi.org/10.1162/EVCO_a_00075.

Huang, C.L., Hsu, N.S., Wei, C.C., Luo, W.J., 2015. Optimal spatial design of capacity and quantity of rainwater harvesting systems for urban flood mitigation. Water (Switzerland) 7, 5173–5202. https://doi.org/10.3390/w7095173.

Huang, J.J., Xiao, M., Li, Y., Yan, R., Zhang, Q., Sun, Y., Zhao, T., 2022. The optimization of Low Impact Development placement considering life cycle cost using Genetic Algorithm. J. Environ. Manag. 309, 114700 https://doi.org/10.1016/j.jenvman.2022.114700.

Huang, Y., Tian, Z., Ke, Q., Liu, J., Irannezhad, M., Fan, D., Hou, M., Sun, L., 2020. Nature-based solutions for urban pluvial flood risk management. WIREs Water 7, e1421. https://doi.org/10.1002/wat2.1421.

Inselberg, A., 2009. Parallel Coordinates: Visual Multidimensional Geometry and its Applications, Parallel Coordinates: Visual Multidimensional Geometry and its Applications. Springer New York, New York, NY. https://doi.org/10.1007/978-0-387-68628-8.

Kaplan, S., 1995. The restorative benefits of nature: toward an integrative framework. J. Environ. Psychol. 15, 169–182. https://doi.org/10.1016/0272-4944(95)90001-2.

Koc, K., Ekmekcioğlu, Ö., Özger, M., 2021. An integrated framework for the comprehensive evaluation of low impact development strategies. J. Environ. Manag. 294, 113023 https://doi.org/10.1016/j.jenvman.2021.113023.

Li, S., Wang, Z., Wu, X., Zeng, Z., Shen, P., Lai, C., 2022. A novel spatial optimization approach for the cost-effectiveness improvement of LID practices based on SWMM-FTC. J. Environ. Manag. 307, 114574 https://doi.org/10.1016/j.jenvman.2022.114574.

Li, X., Zhang, K., Gu, P., Feng, H., Yin, Y., Chen, W., Cheng, B., 2021. Changes in precipitation extremes in the Yangtze River Basin during 1960–2019 and the association with global warming, ENSO, and local effects. Sci. Total Environ. 760, 144244 https://doi.org/10.1016/j.scitotenv.2020.144244.

Liu, L., Li, X., Xia, G., Jin, J., Chen, G., 2016. Spatial fuzzy clustering approach to characterize flood risk in urban storm water drainage systems. Nat. Hazards 83, 1469–1483. https://doi.org/10.1007/s11069-016-2371-4.

MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, pp. 281–296.

Macro, K., Matott, L.S., Rabideau, A., Ghodsi, S.H., Zhu, Z., 2019. OSTRICH-SWMM: a new multi-objective optimization tool for green infrastructure planning with SWMM. Environ. Model. Software 113, 42–47. https://doi.org/10.1016/j.envsoft.2018.12.004.

Maier, H.R., Razavi, S., Kapelan, Z., Matott, L.S., Kasprzyk, J., Tolson, B.A., 2019. Introductory overview: optimization using evolutionary algorithms and other metaheuristics. Environ. Model. Software 114, 195–213. https://doi.org/10.1016/j.envsoft.2018.11.018.

Mani, M., Bozorg-Haddad, O., Loáiciga, H.A., 2019. A new framework for the optimal management of urban runoff with low-impact development stormwater control measures considering service-performance reduction. J. Hydroinf. 21, 727–744. https://doi.org/10.2166/hydro.2019.126.

Matteo, G., Andrea, C., Francesca, P., Emanuele, M., M., R.P., 2016. Curses, tradeoffs, and scalable management: advancing evolutionary multiobjective direct policy search to improve water reservoir operations. J. Water Resour. Plann. Manag. 142, 4015050 https://doi.org/10.1061/(ASCE)WR.1943-5452.0000570.

Mell, I.C., 2010. Green Infrastructure: Concepts, Perceptions and its Use in Spatial Planning. Newcastle University.

Meza, J.C., Oliva, R.A., 2003. An object oriented library to manage the collection of Schittkowski test problems for nonlinear optimization. In: Proceedings of the 2003 Conference on Diversity in Computing. TAPIA '03. https://doi.org/10.1145/948542.948544.

Mitra, P., Uma Shankar, B., Pal, S.K., 2004. Segmentation of multispectral remote sensing images using active support vector machines. Pattern Recogn. Lett. https://doi.org/10.1016/j.patrec.2004.03.004.

Muhammed, K., Farmani, R., Behzadian, K., Diao, K., Butler, D., 2017. Optimal rehabilitation of water distribution systems using a cluster-based technique. J. Water Resour. Plann. Manag. 143, 04017022 https://doi.org/10.1061/(ASCE)WR.1943-5452.0000770.

Nehrke, S.M., Roesner, L.A., 2004. Effects of design practice for flood control and best management practices on the flow-frequency curve. J. Water Resour. Plann. Manag. 130, 131–139. https://doi.org/10.1061/(ASCE)0733-9496(2004)130:2(131).

Nehrke, S.M., Roesner, L.A., 2002. Effect of detention and BMPs on flow frequency of runoff. In: Linking Stormwater BMP Designs and Performance to Receiving Water Impact Mitigation. American Society of Civil Engineers, Reston, VA, pp. 254–265. https://doi.org/10.1061/40602(263)18.

Ng, H.P., Ong, S.H., Foong, K.W.C., Goh, P.S., Nowinski, W.L., 2006. Medical image segmentation using k-means clustering and improved watershed algorithm. In: Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation.

Reynoso-Meza, G., Garcia-Nieto, S., Sanchis, J., Blasco, F.X., 2013. Controller tuning by means of multi-objective optimization algorithms: a global tuning framework. IEEE Trans. Control Syst. Technol. https://doi.org/10.1109/TCST.2012.2185698.

Rossman, L.A., 2017. Storm Water Management Model Reference Manual - Hydraulics, Environmental Protection. US EPA Office of Research and Development, Washington, DC. https://doi.org/10.1145/948542.948544.

Sambito, M., Di Cristo, C., Freni, G., Leopardi, A., 2020. Optimal water quality sensor positioning in urban drainage systems for illicit intrusion identification. J. Hydroinf. 22, 46–60. https://doi.org/10.2166/hydro.2019.036.

Sánchez-Orgaz, S., Pedemonte, M., Ezzatti, P., Curto-Risso, P.L., Medina, A., Calvo Hernández, A., 2015. Multi-objective optimization of a multi-step solar-driven Brayton plant. Energy Convers. Manag. https://doi.org/10.1016/j.enconman.2015.04.077.

Satour, A., Montrésor, S., Bentahar, M., Boubenider, F., 2020. Wavelet based clustering of acoustic emission hits to characterize damage mechanisms in composites. J. Nondestr. Eval. 39, 1–11. https://doi.org/10.1007/s10921-020-00678-1.

Seyedashraf, O., Bottacin-Busolin, A., Harou, J.J., 2022. A design framework for considering spatial equity in sustainable urban drainage infrastructure. Sustain. Cities Soc., 103960 https://doi.org/10.1016/j.scs.2022.103960.

Seyedashraf, O., Bottacin-Busolin, A., Harou, J.J., 2021a. Many-objective optimization of sustainable drainage systems in urban areas with different surface slopes. Water Resour. Manag. 35, 2449–2464. https://doi.org/10.1007/s11269-021-02840-4.

Seyedashraf, O., Bottacin-Busolin, A., Harou, J.J., 2021b. A disaggregation-emulation approach for optimization of large urban drainage systems. Water Resour. Res. 57, e2020WR029098.

Shojaeizadeh, A., Geza, M., Hogue, T.S., 2021. GIP-SWMM: a new green infrastructure placement tool coupled with SWMM. J. Environ. Manag. 277, 111409 https://doi.org/10.1016/j.jenvman.2020.111409.

Sun, X., Lall, U., Merz, B., Dung, N.V., 2015. Hierarchical Bayesian clustering for nonstationary flood frequency analysis: application to trends of annual maximum flow in Germany. Water Resour. Res. 51, 6586–6601. https://doi.org/10.1002/2015WR017117.

Torres, M.N., Fontecha, J.E., Zhu, Z., Walteros, J.L., Rodríguez, J.P., 2020. A participatory approach based on stochastic optimization for the spatial allocation of Sustainable Urban Drainage Systems for rainwater harvesting. Environ. Model. Software 123, 104532. https://doi.org/10.1016/j.envsoft.2019.104532.

Washington State Department of Ecology, Herrera Environmental Consultants, 2012. PugetSound Stormwater BMP Cost Database. Herrera Environmental Consultants, Inc., Washington, United States.

Woods Ballard, B., Wilson, S., Udale-Clarke, H., Illman, S., Scott, T., Ashley, R., Kellagher, R., 2015. The SUDS Manual, Ciria. CIRIA.

Wright, H., 2011. Understanding green infrastructure: the development of a contested concept in England. Local Environ. 16, 1003–1019. https://doi.org/10.1080/13549839.2011.631993.

Zadeh, L.A., 1965. Fuzzy sets. Inf. Control 8, 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X.

Zatarain Salazar, J., Reed, P.M., Quinn, J.D., Giuliani, M., Castelletti, A., 2017. Balancing exploration, uncertainty and computational demands in many objective reservoir optimization. Adv. Water Resour. 109, 196–210. https://doi.org/10.1016/j.advwatres.2017.09.014.

Zio, E., Bazzo, R., 2011. A clustering procedure for reducing the number of representative solutions in the Pareto Front of multiobjective optimization problems. Eur. J. Oper. Res. 210, 624–634. https://doi.org/10.1016/j.ejor.2010.10.021.