

An evolutionary multi-objective optimization system for earthworks

M. Parente^a, P. Cortez^b, A. Gomes Correia^c

^a Corresponding author. ISISE Institute for Sustainability and Innovation in Structural Engineering / ALGORITMI Research Centre. University of Minho, School of Engineering, Azurém, 4800-058 Guimarães, Portugal (email: map@civil.uminho.pt; tel.: +351 253510200)

^b ALGORITMI Research Centre. University of Minho, School of Engineering, Azurém, 4800-058 Guimarães, Portugal (pcortez@dsi.uminho.pt)

^c ISISE Institute for Sustainability and Innovation in Structural Engineering. University of Minho, School of Engineering, Azurém, 4800-058 Guimarães, Portugal (agc@civil.uminho.pt)

Abstract

Earthworks involve the levelling or shaping of a target area through the moving or processing of the ground surface. Most construction projects require earthworks, which are heavily dependent on mechanical equipment (e.g., excavators, trucks and compactors). Often, earthworks are the most costly and time-consuming component of infrastructure constructions (e.g., road, railway and airports) and current pressure for higher productivity and safety highlights the need to optimize earthworks, which is a nontrivial task. [Most previous attempts at tackling this problem focus on single-objective optimization of partial processes or aspects of earthworks, overlooking the advantages of a multi-objective and global optimization.](#) This work describes a novel optimization system based on an evolutionary multi-objective approach, [capable of globally optimizing several objectives simultaneously and dynamically.](#) The proposed system views an earthwork construction as a production line, where the goal is to optimize resources under two crucial criteria (costs and duration) and focus the evolutionary search ([non-dominated sorting genetic algorithm-II](#)) on compaction allocation, using linear programming to distribute the remaining equipment (e.g., excavators). Several experiments were held using real-world data from a Portuguese construction site, showing that the proposed system is quite competitive when compared with current manual earthwork equipment allocation.

Keywords: earthworks, evolutionary computation, multi-objective optimization, artificial intelligence

1. Introduction

In Civil Engineering, a great majority of construction projects require earthworks activities prior to the construction of any structural element. Earthworks are engineering

processes by which the ground surface in a target area is levelled or shaped through the moving or processing of the geomaterials that comprise it. It usually involves the excavation of these geomaterials, which can then be loaded and hauled to new areas to be spread and compacted into embankments, and may also include intermediate steps, such as material treatment or layer wetting. Nowadays, technical and environmental concerns require that, whenever possible, embankment fronts be built using mostly the material excavated from the construction site itself, in order to take maximum advantage of available materials and avoid the use of other materials brought in from outside borrowing areas. Earthwork tasks are reliant on heavy mechanical equipment, namely excavators (material excavation and loading to transportation equipment), dumper trucks (transportation between excavation and embankment fronts), bulldozers (material spreading so as to allow for compaction) and compactors (Zhang, 2008; Hola & Schabowicz, 2010). This is one of the main reasons that make earthworks to often incur the highest percentage costs and durations in road and railway construction projects, implying an increased importance regarding their automated optimization (Miao et al., 2011).

Under this context, it is essential to optimize all available earthwork resources under two key objectives: cost and duration. Both depend on the availability of equipment and also on equipment allocation throughout the project. This is a nontrivial task due to several reasons. There is a vast number of possible equipment allocation combinations, thus the search space is large. And while there is often high competitiveness, where the pressure is to provide the least possible costs and durations (Moselhi & Alshibani, 2007), contractors and project designers often settle for an allocation solution that is mostly based on their own intuition and experience. Moreover, trade-offs need to be set between cost and duration, since these objectives can also conflict. For instance, a less expensive solution may use less amount of equipment, which in turn will result in higher project durations. Contrariwise, allocating more equipment will increase project costs, but decrease durations. However, a prolonged use of equipment implies higher fuel consumption and maintenance expenses. Furthermore, as an earthwork construction progresses, the equipment must be optimally reallocated in order to advance through successive construction phases, characterizing this as a dynamic multi-objective optimization problem.

Ideally, earthworks should be optimized automatically. Considering the nontrivial characteristics of earthworks optimization (e.g., large search space and conflicting goals), pure conventional Operational Research (e.g., linear programming) and blind search methods

are infeasible. An interesting solution is to adopt metaheuristics, which are flexible optimization methods capable of searching interesting search space regions under a reasonable use of computational resources. Several studies have adopted metaheuristics to earthwork optimization, such as genetic algorithms (Marzouk & Moselhi, 2002; Moselhi & Alshibani, 2007; Xu et al., 2011) and swarm intelligence (Kataria et al., 2005; Miao, Sun et al., 2011; Nassar & Hosny, 2012; Zhang, 2008). However, many of these studies focus on single tasks or partial processes that comprise earthworks, such as excavation and hauling (Edwards & Griffiths, 2000; Kataria et al., 2005; Nassar & Hosny, 2012; Xu et al., 2011), in an attempt to deal with the high complexity of the problem. Therefore, these systems lack the advantages of a global optimization of execution durations and costs throughout all construction phases. In terms of optimization objectives, existent systems tend to be limited to single objective optimization, such as cost (Marzouk & Moselhi, 2002) or duration (Kataria et al., 2005), or attempt to consider both objectives via a weight-based optimization (Zhang, 2008). Although these solutions are considered effective in reducing computation effort requirements, they overlook the advantages of optimizing both objectives simultaneously. Even if it can be looked at as multi-criteria optimization, the weighted-based approach used in (Zhang, 2008) only outputs a single trade-off for a particular weight combination (e.g., 0.8 for first criteria and 0.2 for second). Yet, often there is not a single optimal trade-off solution, but rather a set of trade-offs with conflicting objectives. Thus, a much natural multi-criteria optimization approach is to optimize a Pareto front of solutions, where each solution is called non-dominated, or Pareto optimal, if none of the objectives can be improved in value without worsening the other (Bonissone et al., 2009). In the context of earthwork optimization, all Pareto-optimal solutions are considered equally good and the main choice criteria for selecting one solution over the other is often decided by the project designer based on the construction final deadline and/or budget. Providing a Pareto set of optimal solutions is valuable for the construction designer, as earthworks are inherently a dynamic and thus different cost-duration solutions might become better adjusted to the ever-changing site conditions as construction develops. Moreover, additional criteria could be used to support the final decision, such as environmental aspects, which can be assessed by the determination of carbon emissions in each solution. [Indeed, environmental concerns and sustainability have been a rising trend in the past few years \(Chang & Tsai, 2015; Guerin, 2014; Wang, Duan, Wu, & Yang, 2014\).](#) For this reason, a flexible methodology that allows for an easy adjustment of the optimization objectives and can deal with different trade-offs is clearly

advantageous.

Taking into account that Pareto front multi-optimization requires the tracking of a population of solutions, population based Metaheuristics such as evolutionary multi-objective optimization (EMO), have become natural and popular solutions. Following this trend, this work adopts a popular EMO method, the non-dominated sorting genetic algorithm-II (NSGA-II) algorithm, to tackle the multi-criteria optimization problem associated with resource allocation in earthwork construction. In contrast with previous works, a true multi-objective approach is adopted, encompassing the whole earthwork construction phase and outputting a Pareto set of interesting trade-off solutions. Moreover, a novel representation of solutions is proposed, using first an EMO to allocate compaction equipment and then a linear programming to distributing the remaining equipment (e.g., excavators and trucks). Finally, [the proposed system](#) is validated by experimenting with real-world data from a construction site and compared against conventional manual earthwork design.

This paper is organized as follows. Section 2 introduces the earthworks in the context of an optimization problem, including the definition of an objective function and associated constraints. Then, Section 3 presents the adaptation of the EMO algorithm to the earthwork domain, focusing on its representation, dynamic features and algorithmic flow. Next, Section 4 details the experiments conducted and analyses obtained results when validating the proposed system using real-world earthworks data. Finally, conclusions are summarized in Section 5, which also presents future research directions.

2. Earthwork Optimization

From the optimization point of view, the earthworks process can be perceived as a production line based on resources (mechanical equipment) and dependency relations between sequential tasks. The available resources can be allocated to each task in the production line, ranging from excavation and transportation to spreading and compaction equipment. Depending on the amount and type of equipment allocated and other factors, such as material type, the work rate for each task in the production line can be easily computed, since it corresponds to the sum of the work rate of assigned equipment.

Ideally, the added work rate of the equipment allocated for each task should be as close as possible to that of the equipment allocated for the next task, in order to allow a constant flow of material throughout the production line. On the one hand, this prevents idle

times from incurring on the allocated equipment, in cases where the work rate of the previous task is significantly inferior to that of the succeeding task. On the other hand, in cases where the rate of the previous task is significantly superior to the succeeding task, an excessive flow of material that can ultimately obstruct movement throughout the construction site is averted. Therefore, controlling the work rate in each task within a production line is paramount. Accordingly, the main variable associated with the earthworks optimization problem is the amount of equipment allocated in each task, for each construction phase.

2.1. Problem Definition

In production lines with sequential interdependent jobs, the last job determines the speed at which the whole process progresses. Considering that compaction corresponds to the last job in the earthworks production line (Figure 1), it determines the development rate of the whole construction, thus having a direct influence on project durations. Maximizing the work rate in compaction fronts would correspond to the minimal execution duration solution, provided that there is enough equipment in the remaining tasks to support such allocation. In this point of view, an earthworks construction project is divided into a number of production lines, which correspond to the total number of compaction fronts. To each compaction front corresponds a potential production line and its associated equipment, ranging from excavation to compaction tasks. These production lines can work simultaneously and are independent from each other while progressing towards completion. However, whenever a compaction front is completed, the equipment associated with that production line becomes available once again. At this point, a construction phase is considered completed and a new one ensues, demanding a reallocation of the newly available equipment. Subsequently, any optimization attempt on this type of problem must take these factors into consideration, including task interdependence and variable site conditions over time.

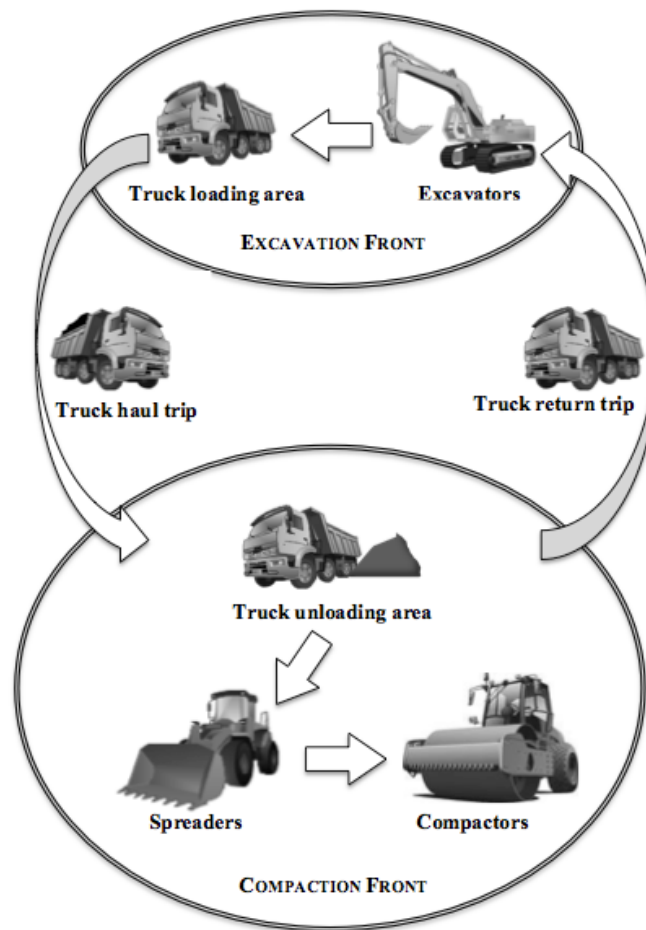


Figure 1. Earthwork equipment aligned with excavation and compaction fronts.

Since resource allocation in earthworks is a time-evolving process (dynamic task), the solution should be time-adaptive as well. More specifically, the allocation of earthwork equipment has to constantly be updated whenever any work front has finished its projected work, hence ending the current construction phase. Consequently, the variables associated with this type of optimization are the amount and type of resources (equipment) to be optimally allocated for each task and also in each construction phase. The algorithmic flow of the general methodology used to solve the allocation and dynamic features of the problem is illustrated in Figure 2.

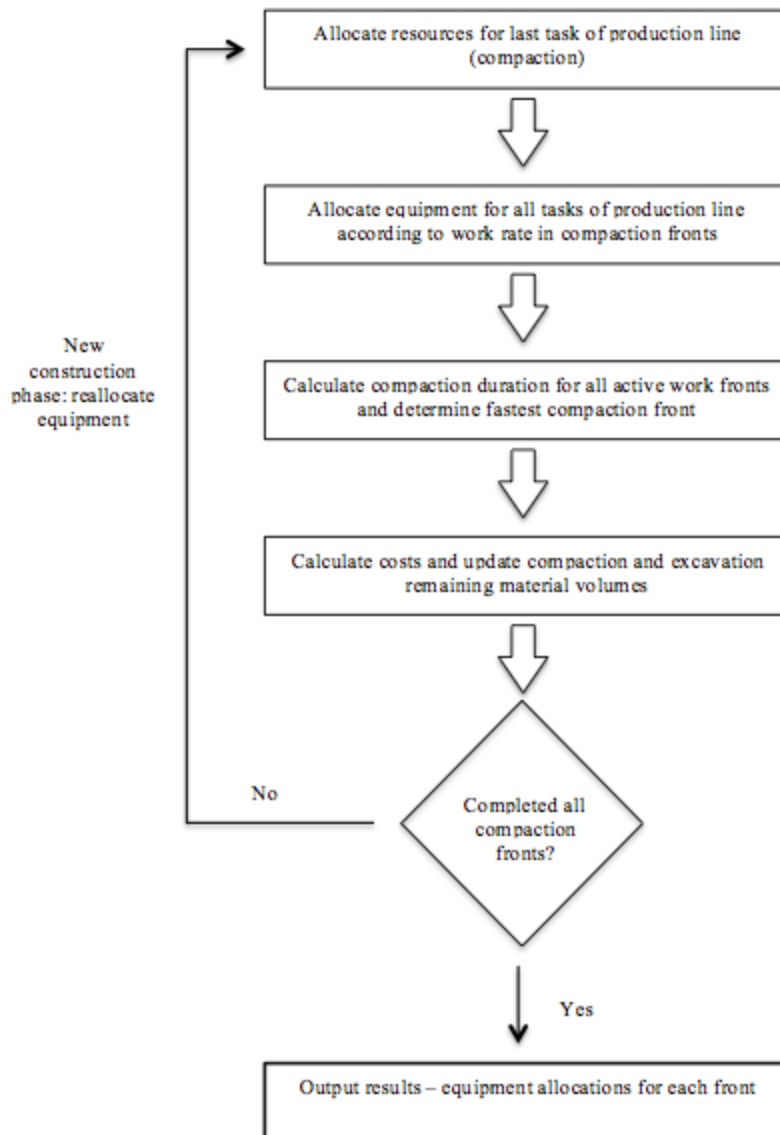


Figure 2. Algorithmic flow of dynamic earthwork resource allocation.

2.2. Objective Function and Constraints

The optimal allocation of earthwork equipment is carried with the main objective of minimizing both total construction times and costs. The ideal solution would allow for the fastest possible excavation, transportation, spreading and compaction of projected volume of geomaterials with minimum expenses. For a single work front, the execution duration, T_s , can be defined by the volume of material being handled (in m^3), V_m , divided by the work rate of the corresponding allocated equipment, Q_p , (in m^3/h), as depicted in Equation 1.

$$T_s = \frac{V_m}{Q_p} \quad (1)$$

However, from the production line point of view, the work rate of each job is not completely independent of that of the adjacent jobs, but rather part of a global work rate of the entire production line. In essence, considering the interdependence of tasks in the production line, the work rate of a production line corresponds to the minimum work rate from between the tasks that comprise it. This occurs because the equipment of one job cannot work at a faster rate than that of the preceding job and, simultaneously, it should not have a higher productivity than the succeeding job. A clear example might be that a group of dumper trucks cannot possibly transport more material than the amount which has been made available by excavators. At the same time, the same dumper trucks should never bring more material to an embankment front than the amount spreaders can handle, since the accumulation of an excessive amount of material at that point can obstruct the remaining equipment. Hence, the duration of the production line work, T_{pl} , corresponds to the volume of material associated with the last job (compaction), V_c , divided by the minimum work rate within the production line, Q_i , as illustrated in Equation 2.

$$T_{pl} = \frac{V_c}{\min(Q_c, Q_s, Q_t, Q_e)} \quad (2)$$

Concurrently, the associated cost can be divided into direct and indirect costs. While the former, $dcost_i$, corresponds to constant expenses, such as equipment rental (when applicable) and manpower costs, the latter, $icost_i$, is related to time dependent costs, such as fuel usage or maintenance cost. In this case, the associated cost can be obtained as shown in Equations 3 (cost of a single piece of equipment, C_s) and 4 (costs per production line, C_{pl}).

$$C_s = \sum icost_i \times T_{pl} + \sum dcost_i \quad (3)$$

$$C_{pl} = \sum C_s \quad (4)$$

Finally, the total duration of the whole earthwork project can be obtained by adding the durations of all non-simultaneous production lines, while total costs are related to the period each equipment is used during the project (this issue is further discussed in Section 3).

It should be noted that the optimization is constrained by the construction deadline, regarding total project duration, and budget, limiting the total costs obtained in the optimal solutions. One additional constraint regarding space restrictions on the construction site can be considered, limiting the maximum amount of equipment simultaneously at work on the same front, when applicable.

2.3. Solution Quality Assessment

As referred in the previous section, total construction time corresponds to the accumulated time for each construction phase. Fundamentally, construction time in each front is a function of equipment productivity and material volume to be handled in that front. In turn, a high amount of factors have influence on the productivity of earthworks equipment for each case. For instance, the productivity of a compactor allocated to a specific front is a function of its type, the type material that is being compacted, and the conditions under which compaction takes place (e.g., layer thickness, atmospheric conditions).

Similarly to total time, total cost equals to the accumulated costs for each construction phase. The latter is a function of the direct and indirect costs linked to each piece of equipment, as well as the amount of hours these are active. Since indirect costs are time dependant (e.g., the fuel usage depends on the amount of time the equipment is active), execution costs can only be calculated subsequently to the determination of construction time in each phase. The steps followed to determine the objective functions (total construction time and cost) are summarized in Table 1.

As the allocation of equipment dictates the resulting construction time and cost, the usage of equipment to its full potential is paramount. In other words, the allocation of equipment takes into account the minimization of construction time and cost, but also the maximization of equipment efficiency. In turn, by using the equipment to its maximum efficiency, the subsequent allocation solutions will inherently reduce the environmental impact of the construction, for instance reducing carbon emissions.

Table 1. Steps for determination of total cost and duration of equipment allocation solutions

Step	Description	Main associated variables/factors
1	Allocation of compactors to embankment fronts	Available compactors (type and quantity); available embankment fronts
2	Determine individual productivity of the allocated compactors for each case	Compactor type; material type; compaction conditions (e.g., layer thickness, meteorological conditions)
3	Calculate total productivity in each active compaction front	Number of compactors of each front and individual productivity of each compactor
4	Allocate spreading equipment	Total productivity in compaction task for associated embankment front; available spreaders/bulldozers (type and quantity); material type; work conditions
5	Allocate transportation equipment	Minimum productivity in spreading and compaction tasks for associated embankment front; available trucks/dumpers (type and quantity); transportation distance; work conditions
6	Allocate excavation equipment	Minimum productivity in transportation, spreading and compaction tasks for associated embankment front; available excavators (type and quantity), material type; work conditions
7	Calculate compaction duration in each embankment front	Productivity of production line (minimum productivity amongst all tasks in a production line); required material volumes for completing each embankment front
8	Verify fastest production line to complete its work (corresponds to the duration of the current construction phase)	Compaction duration of each production line; total volume of material required to complete each active embankment front
9	Calculate volumes of materials which have been excavated and compacted in each front during current phase	Duration of current construction phase; productivity of each production line; volume of material available/required in each active front
10	Calculate cost according to the used equipment and the duration of current construction phase	Direct and indirect costs of active equipment; duration of current construction phase
11	Verify if all embankment fronts have been completed. If not, initiate new construction phase (step 1), taking into account updated material volumes (calculated in step 9). Otherwise, output accumulated cost and duration.	Available embankment fronts (if initiating new construction phase); individual cost and duration for each construction phase (if outputting results)

In this context, and bearing in mind that earthwork construction can be interpreted as a series of production lines, global productivity will be at its highest rate when the productivity of the last task in these production lines (i.e., compaction task) is maximized. Given this premise, the allocation of equipment is firstly carried out for the compaction task, and then for each preceding task, as described in steps 1-6 of Table 1. Additionally, in order to guarantee maximum equipment efficiency, the allocation of the other tasks is performed in function of the productivity verified in embankment fronts.

For example, consider an initial construction phase where three compactors were distributed between two fronts. Lets assume the compactors, c_i , are all of the same type, and the materials, as well as compaction conditions, in front 1 and 2 are the same, resulting in a compaction productivity of $400 \text{ m}^3/\text{h}$ per compactor. This means that total productivity in front 1 would be $400 \text{ m}^3/\text{h}$, while front 2 would proceed at a rate of $800 \text{ m}^3/\text{h}$. As such, the enough spreaders (e.g., bulldozers) must be allocated for each front so that the total productivity is as close as possible (equal or higher) to $400 \text{ m}^3/\text{h}$ in front 1 and $800 \text{ m}^3/\text{h}$. That could mean that a higher amount of spreading equipment can be necessary for front 2 then for front 1, depending on the available spreader characteristics, their productivity when handling the material type, and the work conditions they are subjected to (e.g., atmospheric conditions). As previously stated, should the available spreaders not be enough to maintain the necessary work rate in one of the fronts, then the work rate of the compactors will be limited by the maximum productivity that those spreaders can achieve. The same allocation methodology is used for transportation and excavation equipment until each active embankment team is associated with a production line that can support its productivity. The methodology used for this allocation is described in Section 3.

Having assembled the resulting production lines, the compaction time of each embankment front can be determined by applying Equation 2, as described in Section 2.2. Following the previous example, assuming that there is the necessary equipment to guarantee that the compactors in both fronts are working at their full potential (400 and $800 \text{ m}^3/\text{h}$ for front 1 and 2, respectively), and if the required volume of material to complete both embankments is 8000 m^3 , then the total duration of compaction should be 20 hours for front 1 and 10 hours for front 2. After the first 10 hours, the equipment associated with the production line of embankment front 1 will become idle, which means a new allocation must be performed. As such, construction phase 1 will be considered complete after 10 hours, and the compaction of all other fronts will be interrupted at that time, since a new optimal distribution of equipment can result in a reallocation of the equipment in that production line. The reason for redistribution all equipment after each construction phase is over is discussed in more detail in Section 3. As such, as construction phase 2 begins, the required volumes for completing embankment fronts 1 and 2 will be 0 m^3 (completed) and 4000 m^3 , respectively. These correspond to steps 7-9 in Table 1.

Finally, having the knowledge of the amount of hours each equipment has been active during construction phase 1 (in this case, 10 hours), it is possible to determine the time-

dependent cost (indirect costs) for each active piece of equipment. By adding the result to the direct costs of active equipment (Equation 3), the total cost for each piece of equipment can be calculated. The total cost for the current construction phase will correspond to the sum of the costs associated with of the active equipment (step 10 in Table 1).

Although a construction phase is considered to end as soon as a compaction front is completed, each solution evaluated by the optimization algorithm is only complete when all fronts have been compacted. As such, this process is repeated for each construction phase, calculating the associated time and cost until all fronts have been compacted (step 11 in Table 1). In this case, since at least one embankment front is still not completed (embankment front 2), then at least one more construction phase will be necessary to complete the process. Subsequently, a new construction phase will begin from step 1 of Table 1, in which the required volume for embankment front 2 is now 4000 m³, whereas front 1 no longer is targeted for compactor allocation. Ultimately, the total time and costs are determined in the end of the process by adding the associated values for each construction phase. Section 3 further details some aspects of this process, framing it in the context of a chromosomal representation adopted for an EMO algorithm.

3. Evolutionary Multi-objective Optimization of Earthworks

Different information technologies have been used for earthworks optimization, ranging from data mining techniques (Chapman et al., 2000; Fayyad et al., 1996), such as artificial neural networks (Hola & Schabowicz, 2010; Marques et al., 2008; Schabowicz & Hoła, 2008; Shi, 1999; Tam et al., 2002), multiple regressions (Edwards & Griffiths, 2000) and other techniques (Chou & Tseng, 2011; Liao, Chu, & Hsiao, 2012), to metaheuristics, such as genetic algorithms (Burdett & Kozan, 2014; Marzouk & Moselhi, 2002; Moselhi & Alshibani, 2007; Xu et al., 2011), swarm intelligence (Jang & Topal, 2014; Kataria et al., 2005; Miao et al., 2011; Nassar & Hosny, 2012; Shishvan & Sattarvand, 2015; Zhang, 2008) and other heuristic algorithms (Madow & Pérez-de-la-Cruz, 2004). The data mining based methods tend to focus on the prediction of equipment work rates rather than on direct resource allocation, which is directly approached using metaheuristics and other heuristics.

The work presented in this paper is based on an EMO. Evolutionary computation methods, including genetic algorithms and EMO, work by maintaining a population of individuals (potential solutions), where a chromosome denotes individual data representation

of a solution and gene is a value position in such representation. And designing the chromosome is a key element when adopting evolutionary approaches, as it defines the search space of the problem.

The majority of previous works that adopt metaheuristics only address one or two sequential steps of the earthwork construction (Edwards & Griffiths, 2000; Nassar & Hosny, 2012; Xu et al., 2011). Since we approach the whole earthwork process, a novel solution representation is proposed in this framework, where the individuals represent potential equipment allocations, which includes the front to which each piece of equipment is allocated. The key idea is to simplify solution representation by using domain knowledge and focusing solely on the optimal distribution of compaction equipment (the last task of the production line). The equipment for the other tasks, namely spreading, transportation and excavation, is then distributed according to the initial allocation of the compaction equipment (known as rollers), depending on the sum of work rates in each compaction front. Thus, for each construction phase, the solution is composed of a sequence of C integer genes: $g_1 g_2 g_3 \dots g_C$, where g_i denotes the position of the i -th compactor (or roller) in terms of its compaction front and C represents the total number of compactors. Genes can take any integer value from 0 to the maximum number of target compaction fronts, F (Figure 3). This representation aims firstly to allocate each roller c_i to a compaction front $\in \{0, 1, 2, \dots, F\}$ (0 means that the roller is not allocated), followed by a validation of each solution acquired this way (and that involves a repair strategy). The whole individual (or chromosome) includes all construction phase gene sequences, thus the total number of genes corresponds to the number of available compactors times the number of necessary construction phases: $C \times F$. In the particular case exemplified in Figure 3, there are $C=3$ rollers and $F=3$ compaction fronts, thus individuals are represented using nine genes.

During fitness evaluation, gene values of 0 correspond to withholding the allocation of the specific roller to any front, remaining unused for the current construction front. This option is relevant since it allows the EMO to discover lower duration solutions and also to deal with (or discard) cases in which the available equipment plant for excavators, transporters and spreaders is not enough to support the allocation of all available rollers. Should the latter case be verified, and taking into account that the total rate of a production line corresponds to the minimum work rate obtained among the tasks comprised in it, the resulting solutions will have either infinite or very high durations with simultaneously high costs. This will cause the EMO to discard these solutions as non-optimal in early stages.

c_1	c_2	\dots	c_C	c_1	\dots	c_C	\dots	c_1	\dots	c_C
g_1	g_2	\dots	g_C	g_{C+1}	\dots	g_{2C}	\dots	$g_{C \cdot F - C + 1}$	\dots	$g_{C \cdot F}$
Phase 1				Phase 2			\dots	Phase F		

c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3
2	1	2	1	1	3	3	0	3
Phase 1			Phase 2			Phase 3		

Figure 3. Chromosome representation for a generic distribution (top) and distribution example for a case with $C=3$ rollers and $F=3$ compaction fronts (bottom).

When validating an individual, a repair strategy is used to assure that the solution is feasible. Such repair strategy involves three ordered steps:

1. **Verification of completed compaction fronts in previous construction phases.** When verified, any rollers allocated to the already completed front are instead allocated to the next active front.
2. **Verification of maximum number of rollers in each compaction front.** This constraint (when applicable) is set to deal with space restrictions on embankment fronts that exist in many construction sites. When verified, the rollers that are above the recommended limit of the current front are reallocated to the next active front according to the rule $g_i = (g_{i+1}) \bmod (F+1)$, if possible, or otherwise assigned a value of 0 (not allocated).
3. **Verification that at least one front must be compacted per construction phase.** This verification is related with specific cases when the EMO generates a solution with a value of 0 to all roller allocations. In such cases, one gene is assigned with a value associated with one of the active fronts.

The order of repair strategies takes into account the prevention of harmful interactions between the three verification steps. As an example, should a roller be reallocated to a front in which the roller limit has already been exceeded as a result of the original front being

completed (repair #1), it will be reallocated once more either to a new front or to none (repair #2). Furthermore, the nature of genetic operators, namely crossover and mutation, guarantees the diversity of solutions, preventing any convergence to local optima as a result of the used repair strategies and assuring that a Pareto-optimal curve is found. It should be noted that these repair strategies were designed to be as least “invasive” as possible, in order to avoid influencing or hindering the algorithm convergence.

After assigning the compactors, the allocation of the remaining equipment plant (excavators, transporters and spreaders) is then carried out by means of linear programming (LP) models. As suggested in (Liu & Lu, 2014), LP models are effective in the optimization of partial or secondary aspects of the process, in which less computation power is required. In this work, to each task and to each equipment type is associated a separate LP model that targets the work rate of the last task in the production line. These were executed as standard LP models, where the main objective function is the minimization of equipment costs (Equations 3 and 4). The associated constraints guarantee that the work rates of the allocated equipment for each task, $Q_{s,i}$, are at least equivalent to that of the last task in the production line (compaction) when possible, or otherwise as close as possible, given the maximum amount of available equipment of each kind. The intention behind this type of distribution is to use the least expensive combination of available equipment that allows the succeeding task to operate at 100% work rate, by matching each task to the work rate of the last task in the production line. Equation 5 shows a generic LP model for any equipment kind, where x_i corresponds to the number of allocated equipment units of a given type, with m_i maximum available units.

$$\begin{aligned}
 & \text{Minimize} && \sum_1^{m_i} C_{s,i} \times x_i \\
 & \text{Subject to} && \sum_1^{m_i} Q_{s,i} \times x_i \geq Q_{total,compaction} \\
 & && x_i \geq m_i
 \end{aligned} \tag{5}$$

In a given construction phase, each production line is independent, but several production lines can be at work at the same time if there is enough equipment to support it. Thus, the number of simultaneous production lines is equal to the number of compaction fronts being compacted at the same time, which, in turn, is a consequence of the initial roller

distribution by the EMO. As such, the maximum number of simultaneous production lines at work is equivalent to the maximum number of rollers. Given the allocation of all equipment for the active production lines, the compaction duration of all fronts is calculated according to Equation 2. From the resulting durations, the compaction front with minimum duration (the fastest compaction front) marks the end of the current construction phase. Accordingly, costs for the current construction phase are determined in function of its duration using Equations 3 and 4.

Each construction phase corresponds to the completion of 1 embankment or compaction front. Whenever an embankment is completed, a new construction phase begins and thus a new equipment reallocation must follow, since the equipment corresponding to the completed production line has become idle/available. However, the construction site conditions may change after each construction phase and limiting the equipment reallocation to the idle equipment would not be a guarantee of optimal distribution. The solution is to treat each construction phase as a completely new optimal allocation of equipment, which must take into account the work already developed in previous phases, including site features and conditions, such as available and required material in excavation or compaction fronts. To achieve this, two memory lists are kept during the determination of construction phase durations and costs, concerning completed fronts and remaining fronts. Both lists are updated at the end of each construction phase. The former includes the fronts that have been completed in previous phases, entailing the duration associated with each, as well as the order of completion. The latter, F_R , includes the actual quantities of material in each compaction front, in function of the duration of the previous construction phase and the work rate of the allocated equipment for each front, as illustrated in Equation 6. The determination of compaction durations in the construction phases (other than the first phase) is always carried out in function of the volumes, $F_{R,i}$, maintained in F_R . Subsequently, progress obtained in all the active compaction fronts is saved at the end of each construction phase, thus being accounted for when determining the compaction durations for the new equipment allocation, in the next phase.

$$F_{R,i} = F_{R,i} - \min(T_{pl,i}) \times Q_i \quad (6)$$

Figure 4 shows the algorithmic flow of the EMO algorithm used to tackle the earthworks optimization problem. The initial population involves the random generation of

front indexes to each roller, representing their allocation, as shown in Figure 3. The fitness function includes the repair methodology described in the previous section, as well as the LP models for the allocation of equipment to the remaining tasks. These allow for the determination of construction phase durations and costs using the formulae discussed in Section 2.2, ultimately leading to the evaluation, in terms of total duration and cost, of each solution. The EMO will generate a new population and repeat this process until the target number of generations is achieved, at each point the Pareto-optimal set is presented, along with the outputs mentioned in Figure 4 for each solution.

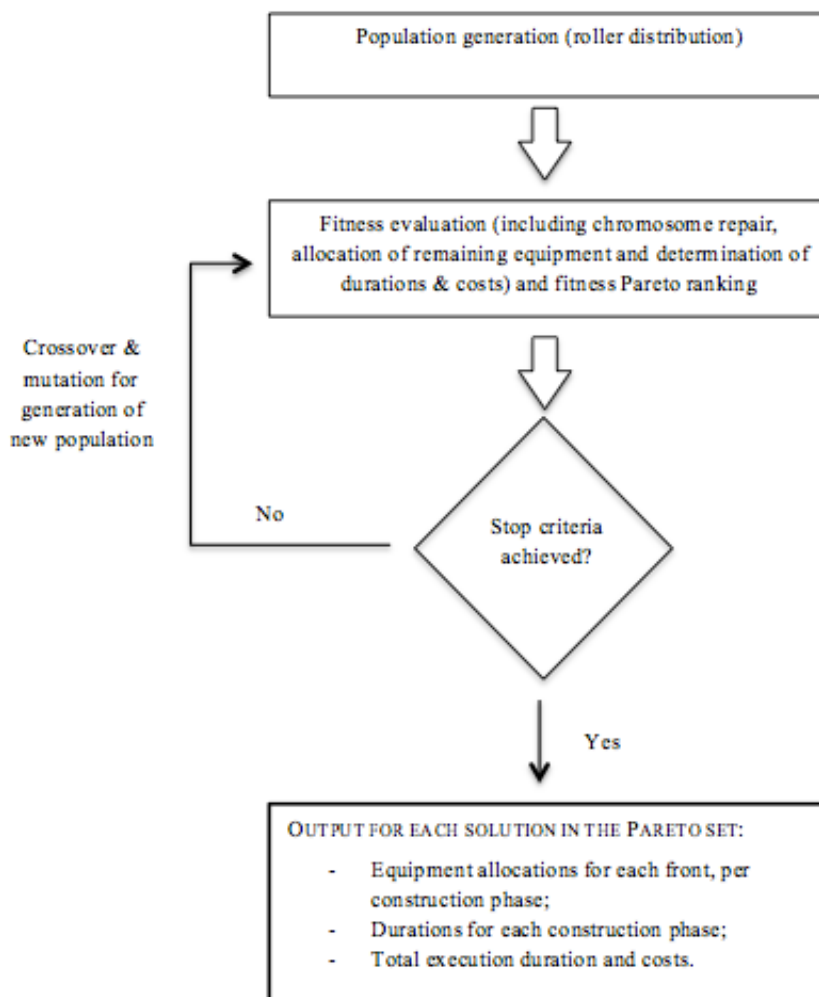


Figure 4. Algorithmic flow of the EMO algorithm.

4. Results

4.1. Real-World Construction Case

To validate the proposed system, we adopt real-world data related with mechanical equipment of a Portuguese road construction site. Originally, the data included the description of several years of earthworks construction, broke down into the daily activities of the available mechanical equipment. The data attributes include date, daily work hours, number and distance of load trips (including daily volumes of transported material) and resource types, as depicted in Table 2.

The data is related with 4 excavation fronts supplying two types of soil (a soil and a soil-rockfill mixture) to 5 embankment fronts. The total volume of transported material is approximately 89,356 m³. The distances from excavation to embankment fronts vary between 100 m to nearly 4,000 m, stressing the need for proper resource management.

Table 2. Example of row values extracted from the available earthwork database.

Date	Work Hrs. (h)	Nr of Load	Associated excavator	Load Zone	Unload Zone	Resource Type	Transp. Volume (m3)
01/6	9			7+850	8+625	Excavator50T	
01/6	8			7+850	8+625	Roller15T	
01/6	10	37	20/871	13+750	12+250	Dumper40T	481
01/6	10	39	20/871	13+750	12+250	Dumper50T	634
01/6	9			13+750	12+250	Tractor40T	

The construction company available equipment included 12 excavators (with sizes ranging from 25 to 75 tons), 22 trucks (either 30 or 40 ton dumper trucks), 8 spreaders (between 20 and 50 ton) and 5 vibrating soil compactors. Both the materials and the compaction equipment were classified according to the *Guide des Terrassements Routiers* (GTR) (SETRA & LCPC, 2000), a well known and broadly used compaction guide. As such, the materials were classified as A1 and C2A1 for the soil and the soil-rockfill mixture, respectively. Accordingly, 4 out of 5 of the vibrating soil compactors were classified as V3, while the remaining one as a V4.

The company designed resource distribution, set taking into account the construction deadlines and budget of the project, resulted in approximately 130 hours of work, spread through 21 days of work (approximately 1 month). All the available equipment was put to

use, with an average value of daily work hours per day equals to 9.88h. Considering this setup, the final cost obtained for these work phases was close to 135,199 €.

The necessary inputs for the optimization system were retrieved from the available database and are summarized in Table 3. Equipment indirect costs and work rates derived from Caterpillar Performance Handbook (Caterpillar Inc., 1998) and Transportation Research Board NCHRP Report 744 (Skolnik, Brooks, & Oman, 2013), as well as the GTR in the case of compaction work rates. Direct costs (including equipment rental and manpower costs) were courtesy of the company.

Table 3. Optimization system inputs

Input	Description	
fa	Embankment front notation and material volume required for completion	Material data
fe	Excavation front notation and material volume available for excavation and transport	
mt	Material type in each excavation front (which will be compacted in embankment fronts after excavation and transport)	
cd	Travel distance matrix from each excavation front to each embankment front	
EE	Available excavation equipment, including type, number of available equipment of each type, work rate and direct and indirect costs	Equipment data
TE	Available transportation equipment, including type, number of available equipment of each type, capacity and direct and indirect costs	
SE	Available spreading equipment, including type, number of available equipment of each type, work rate and direct and indirect costs	
CE	Available compaction equipment, including type, number of available equipment of each type, work rate and direct and indirect costs	

4.2. Computational Experiments

In this work, the equipment available was kept fixed to the one used by the construction company in their conventional allocation design. With a total of $F=5$ production lines working simultaneously and $C=5$ available compactors, the encoded individuals have $5 \times 5 = 25$ genes each, defining the search space for this problem. As the EMO search engine, the Non-dominated Sorting Genetic Algorithm-II (NSGA-II) (Deb, Pratap, Agarwal, & Meyarivan, 2002) was adopted due two main reasons. Firstly, NSGA-II is a popular and standard method for multi-criteria evolutionary optimization. Secondly, NSGA-II is easily

available for a computational use in the R statistical tool (R Development Core Team, 2011) via the package *mco* (Mersmann, Trautmann, Steuer, Bischl, & Deb, 2014), which is the same tool adopted for the development of our integrated optimization system. Moreover, the R tool includes several conventional optimization methods, such as the Linear Programming (LP) method that was used for individual fitness calculation.

The default parameterization of NSGA-II method, as implemented in the R tool, was used, namely: population size of 100, stop after 100 generations, crossover probability of 0.7 and mutation probability of 0.2. The rationale is to focus more on assessing and validating the capabilities of the proposed integrated system when compared with current human design, rather than calibrating the optimization algorithm. Note that in preliminary tests, smaller population sizes (*i.e.*, 20, 30, 40) were explored, but the obtained results were worse than the default population size of 100. Also, the fitness evaluation is computationally costly, as it requires several LP optimizations (for each front), thus a population size much larger than 100 individuals would increase the computational effort. Given that in the *mco* package of the R tool adopts a real value representation for the NSGA-II method, all genes were first rounded to the nearest integer as the first step of the fitness function.

The method was executed with an exclusive access to an Intel Xeon 2.27GHz server under a Linux server. [To get a more robust assessment of the quality of the results, 30 distinct runs of the NSGA-II algorithm were executed. The total computational effort \(considering all 30 runs\) was approximately 254,642.5 seconds \(around 70.73 hours\).](#)

To illustrate the multi-objective convergence, Figure 5 plots the evolution of solutions optimized by the NSGA-II towards the Pareto-optimal front according to cost and duration goals, and during a single run. In the plot, each point denotes a possible solution while line segments are used to join points that belong to the Pareto front. To facilitate the analysis, a colouring scheme is used, ranging from light grey (first generation) to black (last generation). Figure 5 shows that NSGA-II performs an initial fast convergence, with substantial movements of the Pareto front towards the bottom left region, and then the algorithm converges more slowly towards the returned Pareto front, which is non-convex near the bottom left region.

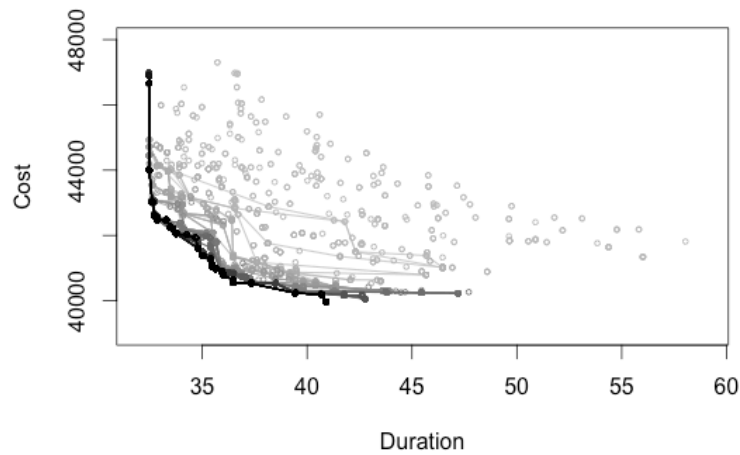


Figure 5. Example of the convergence of NSGA-II algorithm (x -axis in € and y -axis in hours).

The obtained results were also compared with two other optimization algorithms. The purpose of the first comparison is to validate the quality and accuracy of obtained results by performing a comparative study with a different and EMO variant, namely the *S metric selection evolutionary multi-objective algorithm* (SMS-EMOA) (Emmerich, Beume, & Naujoks, 2005). The second comparison is aimed at demonstrating the advantages of multi-objective optimization when compared with a single-objective optimization method. These comparisons are depicted in Figure 6, showing the average Pareto fronts for both EMO algorithms, as well as the average point corresponding to a cost-only optimization solution using single-objective optimization. The average Pareto curves were achieved by performing a vertical averaging procedure (i.e., according to the Duration objective) of the Pareto curves outputted by each run, using the averaging method proposed by Fawcett (2006) for vertical averaging of ROC curves. The graph also includes “H” shape whiskers, on both the lines and the point, that denote the 95% confidence interval bars according to a t-student distribution.

Regarding the comparison with a different EMO variant, the SMS-EMOA was chosen, for being established as a recent and well-known EMO algorithm, which makes use of the hypervolume measure as selection criterion. Furthermore, this algorithm can be implemented under the R tool via a small portion of code that uses functions from the *emoa* package (Mersmann, 2012), facilitating the comparative analysis. Since the aim is to allow a direct comparison between results and performances of both EMO algorithms, the parameters used in SMS-EMOA are the same that were used for the NSGA-II optimization (population=100, stop after 100 generations, crossover probability=0.7, mutation probability=0.2). The analysis of Figure 6 reveals a good consistency between the results

obtained by both EMO algorithms. Nevertheless, NSGA-II slightly outperforms the SMS-EMOA in terms of the resulting Pareto curve results. Furthermore, the total computational effort for the SMS-EMOA to perform all 30 runs was approximately 270,605.1 seconds (around 75.17h), which is also above the 70.73 hours associated with the NSGA-II runs.

The single objective algorithm is implemented in R by adopting the *genalg* package (Willighagen, 2015). The solutions for the single-objective optimization of costs (30 runs with the same optimization parameters as the EMO algorithms) seem to be consistent with the results obtained by the EMO algorithms in terms of the cost reduction goal. However, as previously referred, since the output is a single allocation solution, it lacks the flexibility of the Pareto curve optimization, providing to the decision maker just one cost-duration trade-off solution under one optimization execution, overlooking the advantages of multi-objective optimization. The computational effort for this algorithm (30 runs) was 212,000.9 seconds (58.89 h).

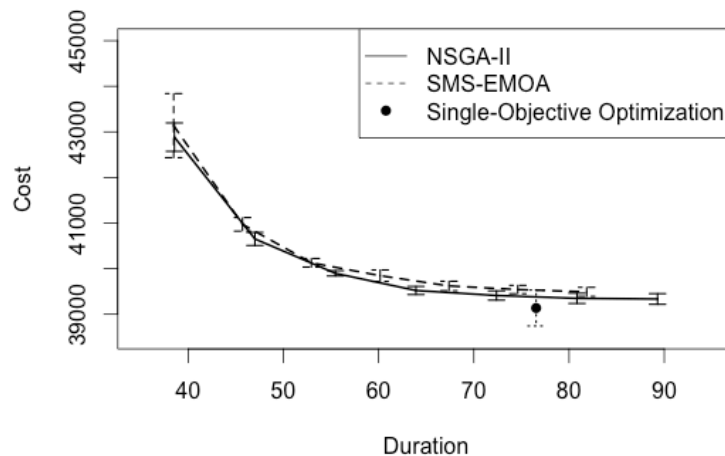


Figure 6. Comparison between results obtained by different optimization methods (x -axis is in hours and y -axis in €)

4.3. Comparison with Manual Allocation

The proposed system results were compared against the original manual allocation performed by the construction company, as shown in Figure 7. In the figure, the left graph shows the average (over all 30 runs) NSGA-II optimized Pareto front, while the right graph compares these results (black line) with the original solution obtained by manual allocation (white point).

When analysing Figure 7, it is clear that NSGA-II performs a substantial improvement (both in terms of cost and duration) when compared with the manual equipment allocation. The system output indicated several potential setups ranging from approximately 38 to 73 hours of construction duration, associated with approximate costs of 40,000 € to 43,000 €, respectively. This corresponds to a reduction of around 55% in duration and 70% in cost, if compared to the duration of 127 h and cost of 135,199 € that was obtained in the original allocation. The original human based allocation is far from the optimal solution, which can be explained by the lack of an efficient automated optimization and also by unforeseen delays.

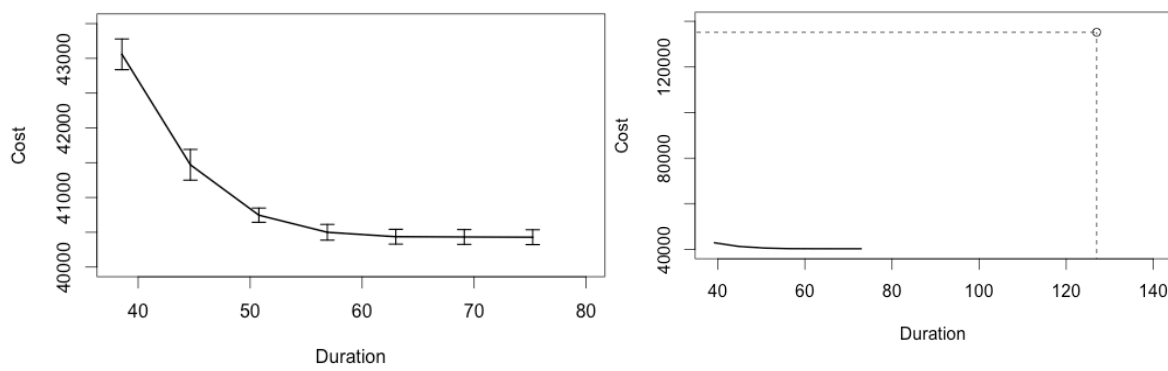


Figure 7. Optimization results ($k=0.75$) in terms of: left graph - vertically averaged Pareto front; and right - comparison between the optimized Pareto front (black curve) and the human based allocation solution (circle point) (in both graphs x -axis is in hours and y -axis in €).

Regarding the lack of efficient automation, the original setup performs bottlenecks in the production lines in one of the tasks preceding compaction, resulting in several occurrences of idle equipment. To illustrate this, two distinct production lines are further analysed in Table 4. It is easy to infer that, for both production lines the work rates in each task of the original distribution setup are not homogeneous, as opposed to the work rates of the optimized solution. In these cases, the whole production line is limited by the work rate of excavators in the original setup, which means that the other tasks have to wait for material to be excavated in order to allow for its transport, spreading and finally compaction. This incurs in equipment idle time while waiting for material to be ready for handling, which represents wastes in terms of resources (since these do not work at full efficiency) and fuel (contributing to unnecessary costs), as well as an increase on unnecessary carbon emissions. Consequently, the total work rate of these production lines cannot be considered superior to that of the minimum work rate obtained in the production line tasks, in this case excavation ($394 \text{ m}^3/\text{h}$ in production line 1

and 540 m³/h in production line 2). In contrast, the work rates obtained in the proposed optimized solutions for each task that comprises the production line are as homogeneous as possible, given the available equipment. As such, a constant flow of material throughout tasks can be achieved, using the allocated resources to their full potential and efficiency.

Table 4. Comparison between original and optimized ($k=0.75$) setups for both cases

Original setup 1		Original setup 2	
Average distance between excavation and compaction fronts: 175 m		Average distance between excavation and compaction fronts: 500 m	
Hauled material volume: 29,753 m ³		Hauled material volume: 10,647 m ³	
<i>Equipment distribution and work rate</i>		<i>Equipment distribution and work rate</i>	
1 Excavator (50T):	394 m ³ /h	1 Excavator (75T):	540 m ³ /h
2 Dumper trucks (30T, 40T):	2,960 m ³ /h	3 Dumper trucks (40T):	1,280 m ³ /h
1 Spreader (20T):	413 m ³ /h	1 Spreader (40T):	675 m ³ /h
1 Vibratory roller (15T):	614 m ³ /h	1 Vibratory roller (19T):	683 m ³ /h
Minimal production line work rate: 394 m ³ /h (excavation)		Minimal production line work rate: 540 m ³ /h (excavation)	
Duration: 75.5 h		Duration: 19.7 h	
Cost: 24,462 €		Cost: 7,996 €	
Optimized setup 1		Optimized setup 2	
Average distance between excavation and compaction fronts: 175 m		Average distance between excavation and compaction fronts: 500 m	
Hauled material volume: 29753 m ³		Hauled material volume: 10647 m ³	
<i>Equipment distribution and work rate</i>		<i>Equipment distribution and work rate</i>	
2 Excavator (75T):	1,080 m ³ /h	2 Excavators (25T, 75T):	743 m ³ /h
2 Tipper trucks:	1,600 m ³ /h	2 Dumper trucks (30T):	880 m ³ /h
2 Spreaders (20T, 50T):	1,239 m ³ /h	1 Spreader (50T):	820 m ³ /h
1 Vibratory roller (19T):	1,055 m ³ /h	1 Vibratory roller (19T):	683 m ³ /h
Minimal production line work rate: 1,080 m ³ /h (compaction)		Minimal production line work rate: 683 m ³ /h (compaction)	
Duration: 28.2 h		Duration: 15.6 h	
Cost: 12,718 €		Cost: 5,740 €	

It is noteworthy to emphasize that, besides optimizing the whole allocation in terms of costs and durations, the developed system is expected to always keep the allocated equipment working at full efficiency. It achieves this by focusing on the minimization of equipment idle time, which will also result in minimization of unnecessary carbon emissions. This last aspect is very challenging to accomplish by conventional design methodologies.

Comparing the original setup with its optimized counterpart for production line 1, the human based allocation solution features a clear excess of work capacity regarding transportation equipment that is not contributing for its progress, as it is limited by the work rate of the excavation team. In order to counter this, the optimization system allocated smaller trucks (lower capacity, lower fuel consumption and, thus, lower operation costs) to fulfil this role instead, while investing its resources more heavily on the excavation, spreading and compaction teams. As a result, the optimized setup for this case is as homogeneous as possible in terms of work rate between tasks (taking into consideration the available equipment), resulting in a decrease of 50% in both duration and cost for this production line. Although the proposed optimized solution might allocate more equipment with a higher productivity, the reduction in duration is significant enough to reduce the operational costs, ultimately resulting in a reduction of both factors.

In other cases, as one can see in the setups for production line 2, which features different fronts and material types, the setup achieved by the optimization module consists of slight differences when compared with the original allocation. The subtle changes to the class of equipment used in some tasks (e.g., spreading), as well as the removal of unnecessary equipment from others (e.g., transportation) allows a greater homogeneity of work rates, resulting in nearly 20% reduction in duration and 30% in cost. Moreover, the equipment that is dismissed from this production line is then available to be used in different work fronts as deemed necessary.

The obtained results emphasize the importance of using intelligent computational tools for optimizing earthworks. In particular, it was shown how conventional human design allocation methodologies can be relatively counter-productive in some situations. During optimization, an efficiency factor of $k=0.75$ was assumed for compactors, as well as other equipment types. The efficiency factor is related to the amount of time that mechanical equipment spends in actual production. According to the technical guides used during the development of this work, namely GTR (SETRA & LCPC, 2000) and Caterpillar Performance Handbook (Caterpillar Inc.,1998), actual “on-the-job” productivity is influenced

by factors such as operator skill, personal delays, job layout and other delays. Since the goal is to maximize resource usage and minimize idle times, it makes sense to use a k value of 0.75, considering that this is the maximum recommended value by the GTR for compaction activities.

Unforeseen delays are due to unpredictable situations that can occur in a real environment, such as presence of bad material or equipment malfunction. The huge difference between NSGA-II results and the manual design might be a consequence of a lower efficiency occurred in the real construction project. To attest the competitiveness of the proposed system, additional NSGA-II experiments were conducted assuming a much lower efficiency factor of $k=0.3$ (Figure 8). When compared with previous experiments of $k=0.75$, the new $k=0.3$ NSGA-II Pareto front solutions require higher costs and duration. In particular, duration now ranges from 105 h to 155 h, a range on which the manual allocation solution falls on (as shown in the right of Figure 8). However, even when assuming a low efficient factor, NSGA-II still outperforms the manual solution, returning a Pareto front that is more interesting and that includes a trade-off point that is much less costly for the same duration.

These new experiments show the versatility of the proposed system in respect to unpredictable events. If needed, the proposed system can be rerun in order to include updates of the site conditions and new restraints, such as altering the available equipment and/or productivity rates.

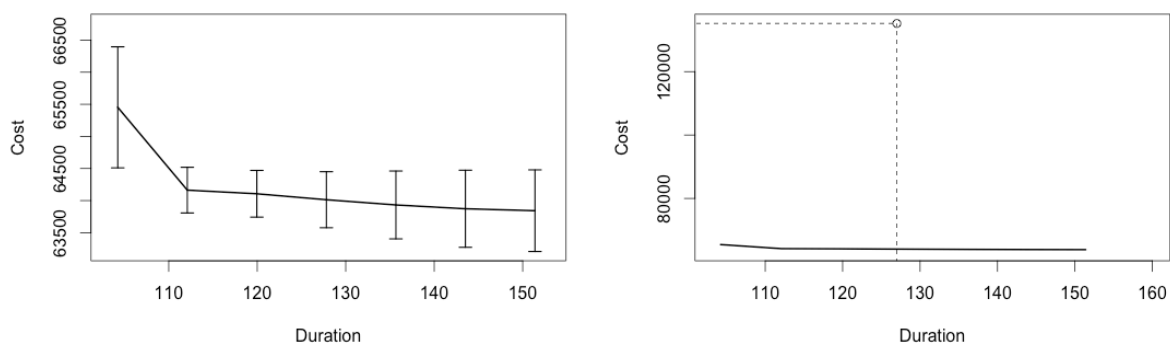


Figure 8. Optimization results ($k=0.3$) in terms of: left graph - vertically averaged Pareto front; and right - comparison between the optimized Pareto front (black curve) and the human based allocation solution (circle point) (in both graphs x -axis is in hours and y -axis in €).

5. Conclusions

Most civil construction structures (e.g., road, railway and airport construction) require earthworks activities, which involve levelling the ground surface and preparing the required foundation conditions. Under this context, the optimization of earthwork is a key factor for construction competitiveness, since earthworks account for a high percentage of the overall construction costs and duration. Conventional manual design, often based on experience and intuition, is limited. Moreover, current intelligent automated applications often target single tasks or partial processes that comprise earthworks and do not simultaneously optimize both cost and duration goals.

This work presents a novel evolutionary multi-objective optimization (EMO) approach that addresses the whole earthwork construction phase and optimizes both cost and duration objectives. It is supported by an original representation of earthworks as production lines, a Pareto approach for dealing with conflicting objectives, and the integration of different optimization methods. The approach has been directed towards a practical and flexible expert intelligent system that can be used in any construction that includes large volumes of earthworks. The potential of the developed system is shown by its validation using real-world data from a Portuguese construction site, while identifying some significant limitations of conventional manual earthworks design. In fact, the obtained results show that the EMO system is quite competitive when compared with the conventional design and that it can be easily adapted to dynamic changes that are inherent to earthworks constructions.

The development of this system opens several new insights, such as: (1) further experimentation with different optimization objectives, in order to keep up with current and future trends, such as environmental impact during construction (e.g., carbon emissions). For example, there has been some development regarding methodologies for controlling and reducing carbon emissions during construction, and such data could be used to test and validate its use as a minimization objective; (2) the optimization methodology regarding the production lines can be enhanced by taking into account other aspects, such as loading and dumping configurations; (3) the system can also profit from focusing on the tuning of the optimization parameters, as well as exploring other EMO variants (e.g., SPEA-2, IBEA, MOEA/D).

Finally, the developed optimization system is expected to be enhanced with data mining and geographic information systems (GIS) technologies, in order to integrate a more

complex system. Data mining can be used to support parameter estimation, such as equipment productivity in a given situation, when the actual data is unavailable (i.e., early design phase), while GIS can provide a visual support for the designer, both in terms of data input (i.e., positioning of excavation and compaction fronts, or determination of shortest routes between fronts) and in terms of visualization of results (i.e., graphically representing the allocation of equipment throughout work fronts).

ACKNOWLEDGEMENTS

The authors wish to thank FCT for the financial support under the doctoral Grant SFRH/BD/71501/2010, as well as the construction company that kindly provided the real-world data. Also, we wish to thank Olaf Mersmann for kindly providing the R code for the SMS-EMOA algorithm.

REFERENCES

- Bonissone, P., Subbu, R., & Lizzi, L. (2009). Multi Criteria Decision Making (MCDM): A Framework for Research and Applications. *IEEE Computational Intelligence Magazine*, 4(3), 48–61.
- Burdett, R. L., & Kozan, E. (2014). An integrated approach for earthwork allocation, sequencing and routing. *European Journal of Operational Research*, 238(3), 741–759. doi:10.1016/j.ejor.2014.04.036
- Caterpillar Inc. (2011). *Caterpillar Performance Handbook Edition 29*.
- Chang, A. S., & Tsai, C. Y. (2015). Sustainable design indicators: Roadway project as an example. *Ecological Indicators*, 53, 137–143. doi:10.1016/j.ecolind.2015.01.036
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0 Step-by-step data mining guide*. Retrieved from <http://www.citeulike.org/group/1598/article/1025172>
- Chou, J.-S., & Tseng, H.-C. (2011). Establishing expert system for prediction based on the project-oriented data warehouse. *Expert Systems with Applications*, 38(1), 640–651. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0957417410006329>
- Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2), 182–197. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=996017

- Edwards, D. J., & Griffiths, I. J. (2000). Artificial intelligence approach to calculation of hydraulic excavator cycle time and output. *Mining Technology*, 109(1), 23–29.
- Emmerich, M., Beume, N., & Naujoks, B. (2005). An EMO algorithm using the hypervolume measure as selection criterion. *Evolutionary Multi-Criterion Optimization*, 62–76. doi:10.1007/978-3-540-31880-4_5
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874. doi:10.1016/j.patrec.2005.10.010
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *American Association for Artificial Intelligence*, 17(3), 1–18.
- Guerin, T. (2014). Root causes of fluid spills from earthmoving plant and equipment: Implications for reducing environmental and safety impacts. *Engineering Failure Analysis*, 45, 128–141. doi:10.1016/j.engfailanal.2014.06.011
- Hola, B., & Schabowicz, K. (2010). Estimation of earthworks execution time cost by means of artificial neural networks. *Automation in Construction*, 19(5), 570–579. doi:10.1016/j.autcon.2010.02.004
- Jang, H., & Topal, E. (2014). A review of soft computing technology applications in several mining problems. *Applied Soft Computing Journal*, 22, 638–651. doi:10.1016/j.asoc.2014.05.019
- Kataria, S., Samdani, S. A., & Singh, A. K. (2005). Ant Colony Optimization in Earthwork Allocation. *International Conference on Intelligent Systems*, (7), 1–9. Retrieved from http://www.geocities.ws/saurabhsamdani/sourcecodes_files/icis-paper.pdf
- Liao, S., Chu, P.-H., & Hsiao, P.-Y. (2012). Data mining techniques and applications – A decade review from 2000 to 2011. *Expert Systems with Applications*, 39(12), 11303–11311. doi:10.1016/j.eswa.2012.02.063
- Liu, C., & Lu, M. (2014). Optimizing Earthmoving Job Planning Based on Evaluation of Temporary Haul Road Networks Design for Mass Earthworks Projects. *Journal of Construction Engineering and Management*. doi:http://dx.doi.org/10.1061/(ASCE)CO.1943-7862.0000940
- Madow, L., & Pérez-de-la-Cruz, J.-L. (2004). Sindi: an intelligent assistant for highway design. *Expert Systems with Applications*, 27(4), 635–644. doi:10.1016/j.eswa.2004.06.005
- Marques, R., Gomes Correia, A., & Cortez, P. (2008). Data Mining Applied to Compaction of Geomaterials. In *Eight International Conference on the Bearing Capacity of Roads, Railways and Airfields*. Montreal, Canada.
- Marzouk, M., & Moselhi, O. (2002). Selecting Earthmoving Equipment Fleets Using Genetic Algorithms. In E. Yucesan, C.-H. Chen, J. L. Snowdon, & J. M. Charnes (Eds.),

- Proceedings of the 2002 Winter Simulation Conference* (pp. 1789–1796). Montreal, Canada.
- Mersmann, O. (2012). Package “emoa”: Evolutionary Multiobjective Optimization Algorithms. Retrieved from <http://www.statistik.tu-dortmund.de/~olafm/software/emoa/>
- Mersmann, O., Trautmann, H., Steuer, D., Bischl, B., & Deb, K. (2014). Package “mco”: Multiple Criteria Optimization Algorithms and Related Functions. Retrieved from <http://git.p-value.net/p/mco.git>
- Miao, K., Sun, X., & Li, L. (2011). A roadbed earthwork allocation model based on ACO algorithm. *Applied Mechanics and Materials*, 44-47, 3483–3486.
- Moselhi, O., & Alshibani, A. (2007). Crew optimization in planning and control of earthmoving operations using spatial technologies. *Journal of Information Technology in Construction*, 12, 1–17. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.98.1818&rep=rep1&type=pdf>
- Nassar, K., & Hosny, O. (2012). Solving the Least-Cost Route Cut and Fill Sequencing Problem Using Particle Swarm. *Journal of Construction Engineering and Management*, 138(8), 931–942.
- Parente, M., Gomes Correia, A., & Cortez, P. (2014). Artificial Neural Networks Applied to an Earthwork Construction Database. In D. Toll, H. Zhu, A. Osman, W. Coombs, X. Li, & M. Rouainia (Eds.), *Second International Conference on Information Technology in Geo-Engineering* (pp. 200–205). Durham, UK: IOS Press.
- R Development Core Team. (2011). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.r-project.org/>
- Schabowicz, K., & Hoła, B. (2008). Application of artificial neural networks in predicting earthmoving machinery effectiveness ratios. *Archives of Civil and Mechanical Engineering*, 8(4), 73–84. doi:10.1016/S1644-9665(12)60123-X
- SETRA, & LCPC. (2000). *Guide des Terrassements Routiers - Réalisation des Semblais et des Couches de Forme*. Paris, France: Laboratoire Central des Ponts et Chaussées.
- Shi, J. J. (1999). A neural network based system for predicting earthmoving production. *Construction Management and Economics*, 17(4), 463–471.
- Shishvan, M. S., & Sattarvand, J. (2015). Long term production planning of open pit mines by ant colony optimization. *European Journal of Operational Research*, 240(3), 825–836. doi:10.1016/j.ejor.2014.07.040
- Skolnik, J., Brooks, M., & Oman, J. (2013). *NCHRP Report 744 - Fuel Usage Factors in Highway and Bridge Construction*. Washington, D.C.

Tam, C. M., Tong, T., & Tse, S. (2002). Artificial neural networks model for predicting excavator productivity. *Journal of Engineering Construction and Architectural Management*, 9(5-6), 446–452.

Wang, X., Duan, Z., Wu, L., & Yang, D. (2014). Estimation of carbon dioxide emission in highway construction: a case study in southwest region of China. *Journal of Cleaner Production*. doi:10.1016/j.jclepro.2014.10.030

Willighagen, E. (2015). Package “genalg.” Retrieved from <https://github.com/egonw/genalg>

Xu, Y., Wang, L., & Xia, G. (2011). Research on the optimization algorithm for machinery allocation of materials transportation based on evolutionary strategy. *Procedia Engineering*, 15, 4205–4210. doi:10.1016/j.proeng.2011.08.789

Zhang, H. (2008). Multi-objective simulation-optimization for earthmoving operations. *Automation in Construction*, 18(1), 79–86. doi:10.1016/j.autcon.2008.05.002