Universidade do Minho
Escola de Engenharia

Manuel Afonso Parente

Development of an Intelligent Earthwork Optimization System

julho de 2015

Universidade do Minho
Escola de Engenharia

Manuel Afonso Parente

Development of an Intelligent Earthwork Optimization System

Tese de Doutoramento em Engenharia Civil

Trabalho efetuado sob a orientação do
Professor Doutor António Gomes Correia
e do
Professor Doutor Paulo Alexandre Ribeiro Cortez

julho de 2015

STATEMENT OF INTEGRITY

I hereby declare having conducted my thesis with integrity. I confirm that I have not used plagiarism or any form of falsification of results in the process of the thesis elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

University of Minho, 09/07/2015

Full name: Manuel Afonso Parente

Signature: Manuel Afonso Parente

ACKNOWLEGDMENTS

This research work has been developed in the scope of the Doctoral Programme in Civil Engineering, at the University of Minho, Portugal, and supported by the Portuguese Science and Technology Foundation (FCT), under grant No. SFRH/BD/71501/2010. This is gratefully acknowledged.

I would like to express my sincere gratitude to Professor António Gomes Correia, for his knowledge, guidance and patience. The respect I owe him, and have for him, got me through some of the greater challenges in the last few years.

I would also like to thank Professor Paulo Cortez for the mind-expanding knowledge he has transmitted me throughout these years. I have learned something new and valuable in each one of our many discussions.

My thanks to my colleague Doctor Joaquim Tinoco, for providing the technical help and insightful advice that allowed me to overcome several obstacles.

I am grateful to my mother, Maria Ascensão Bastos, for teaching me that the bigger the challenge, the greater the reward. Furthermore, her expertise regarding the English language has been one of the pillars that supported the writing of this work.

I appreciate the time spent with my father, Hernâni Parente, in which he opened my eyes to another side of life, and showing me that developing one's spirit is just as important as training one's mind and body.

Thank you, Aline Santos, for constantly pointing me towards the right path and demonstrating what a brighter attitude can accomplish.

Last but not least, I would like to thank my family and friends for the good times, encouragement and support.

ABSTRACT

Earthworks are often regarded as one of the most costly and time-consuming components of linear infrastructure constructions (e.g., road, railway and airports). Since actual construction requirements originate higher demands for productivity and safety in earthwork constructions, the optimal usage of every resource in these tasks is paramount. The management of resources in an earthwork construction site is, in great part, a function of the allocation of the available equipment, for which there are a vast number of possible equipment allocation combinations. Simultaneously, while there is often high competitiveness, where the pressure is to provide the least possible costs and durations, contractors and project designers often settle for an allocation solution that is mostly based on their own intuition and accumulated experience. This guarantees neither optimal resource usage, nor a solution associated with minimal cost and duration.

The optimal allocation of equipment in earthwork tasks is a complex problem that requires the study of several different aspects, as well as the knowledge of a large number of factors. In fact, earthworks are comprised by a combination of repetitive, sequential, and interdependent activities based on heavy mechanical equipment (i.e., resources), such as excavators, dumper trucks, bulldozers and compactors. In order to optimally allocate the available resources, knowledge regarding their specifications (e.g., capacity, weight, horsepower) and the work conditions to which they will be subjected (e.g., material types, required and available volumes in embankment and excavation fronts, respectively) is essential. This knowledge can be translated into the productivity (i.e., work rate) of each piece of equipment when working under a specific set of conditions. Moreover, since earthwork tasks are inherently sequential and interdependent, the interaction between the allocated equipment must be taken into account. A typical example of this is the need for matching the work rate of an excavator plant with the capacity of a truck plant to haul the excavated material to the embankment fronts.

Given the non-trivial characteristics of the earthwork allocation problem, conventional Operation Research (e.g., linear programming) and blind search methods are infeasible. As such, a potential solution is to adopt metaheuristics – modern optimization methods capable of searching large space regions under a reasonable use of computational resources. While this may address the issue of optimizing such a complex problem, the lack of knowledge regarding optimization parameters under different work conditions, such as equipment productivity, calls for a different approach. Bearing in mind the availability of large databases, including in the earthworks area, that have been gathered in recent years by construction companies, technologies like data mining (DM) come forward as ideal tools for solving this problem. Indeed, the learning capabilities of DM algorithms can be applied to databases embodying the productivity of several equipment types when subjected to different work conditions. The extracted knowledge can then be used to estimate the productivity of the available equipment under similar work conditions. Furthermore, as previously referred, since earthwork tasks include the material hauling from excavation to embankment fronts, it also becomes imperative to analyse and optimize the possible transportation networks. In this context, the use of geographic information systems provides an easy method to study the possible trajectories for

transportation equipment in a construction site, ultimately allowing for a choice of the best paths to improve the workflow.

This work explores the integration of different technologies in order to allow for an optimization of the earthworks process. This is translated in the form of an evolutionary multi-criteria optimization system, capable of searching for the best allocation of the available equipment that minimizes a set of goals (e.g., cost, duration, environmental impact). The results stemming from the application of the system to a case study in a Portuguese earthwork construction site are presented. These comprise the assessment of the system performance, including a comparison between different optimization methods. Furthermore, an analysis regarding the improvement of workflow in the construction site after the implementation of the system is discussed, in the context of several comparisons between original (i.e., obtained by manual design) and optimized allocation solutions. Ultimately, these results illustrate the potential and importance of using this kind of technologies in the management and optimization of earthworks.

RESUMO

Em projetos de construção de infraestruturas de transporte lineares (e.g., estradas, vias férreas e aeroportos), as terraplenagens são geralmente consideradas um dos componentes com custos e tempos de execução mais elevados. Tendo em conta que cada vez mais é exigido um aumento na produtividade e segurança no contexto das construções de terraplenagens, torna-se fulcral a otimização de todas as tarefas relacionadas com este processo. A gestão de recursos num estaleiro de terraplenagens é, em grande parte, função da alocação do equipamento mecânico disponível, para a qual existe um número quase infinito de soluções possíveis em cada caso. Simultaneamente, embora se verifique um alto nível de competitividade nesta área, onde o objetivo é obter custos e durações de execução o mais baixos possíveis, o planeamento das tarefas de terraplenagens é em grande parte baseado na experiência acumulada dos engenheiros e especialistas. Porém, tais métodos não garantem nem uma utilização ótima dos recursos disponíveis, nem uma solução associada ao custo e duração de execução mínimos.

A alocação ótima de equipamento mecânico em tarefas de terraplenagens é um problema complexo que requer o estudo de vários aspectos distintos, assim como o conhecimento de um elevado número de fatores. De facto, estas tarefas são demarcadas por combinações de atividades repetitivas, fortemente baseadas no uso de equipamento mecânico (i.e., recursos), tal como escavadoras, *dumpers*, espalhadores e compactadores. Para que seja possível a sua alocação ótima, é essencial o conhecimento das suas especificações (e.g., capacidade, peso, potência) e das condições a que estão sujeitos durante a sua atividade (e.g., tipos de material, volumes disponíveis em frentes de escavação e necessários em frentes de aterro). Este conhecimento pode ser traduzido na produtividade de cada equipamento quando sujeito a determinadas condições de trabalho. Para além disso, uma vez que as terraplenagens consistem em tarefas inerentemente sequenciais e interdependentes, a interação entre os equipamentos tem de ser tomada em consideração. Um exemplo típico deste aspecto pode ser ilustrado pela necessidade de sincronizar a produtividade de uma equipa de escavadoras com a de uma equipa de *dumpers*, para que seja possível um fluxo constande de escavação e transporte de geomateriais das frentes de escavação para as frentes de aterro.

Tendo em conta as características não triviais do problema de alocação em terraplenagens, os métodos convencionais de procura de soluções, tais como Investigação Operacional (e.g. programação linear) e busca exaustiva são impraticáveis. Assim, uma potencial solução é a adoção de metaheurísticas – métodos de otimização moderna capazes de efetuar a busca de soluções em espaços de procura extensos com níveis de exigência computacional razoáveis. Embora estes métodos sejam práticos para a otimização de problemas de elevado nível de complexidade, como é o caso das terraplenagens, existe ainda a necessidade de abordar o problema relacionado com a escassez de conhecimento de vários parâmetros necessários à otimização, tais como a produtividade dos equipamentos sujeitos a diferentes condições de trabalho. Considerando os recentes avanços da tecnologia e o aumento da prática de recolha de dados, verifica-se a disponibilidade de extensas bases de dados de construção, incluindo na área de terraplenagens. Neste sentido, tecnologias tais como o *data mining* (DM) surgem como ferramentas ideais para abordar esse problema. De fato, as capacidades de aprendizagem dos

algoritmos de DM podem ser aplicadas às bases de dados existentes com informação relativa à produtividade de vários tipos de equipamento sujeitos a diferentes condições de trabalho. Mediante este processo, o conhecimento extraído pode então ser usado em novos casos para estimar a produtividade de equipamentos em condições semelhantes. Adicionalmente, uma vez que as tarefas de terraplenagens incluem o transporte de materiais de frentes de escavação para frentes de aterro, como previamente referido, torna-se ainda imperativa a análise e otimização das potenciais trajetórias de transporte ao longo do estaleiro. Neste contexto, a utilização de sistemas de informação geográficos providencia um método eficaz de estudo e escolha das melhores trajetórias para o equipamento de transporte, melhorando o fluxo de trabalho no estaleiro.

Este trabalho explora a integração de diferentes tecnologias tendo em vista a otimização das tarefas de terraplenagens. Isto concretiza-se sob a forma de um sistema de otimização evolutiva multi-objetivo, capaz de eleger a melhor distribuição dos equipamentos de terraplenagens disponíveis que minimiza um determinado conjunto de objetivos (e.g., custo, duração, impacto ambiental). São apresentados os resultados decorrentes da aplicação do sistema desenvolvido num caso de estudo, associado a um estaleiro de terraplenagens em Portugal. Estes abrangem a avaliação do desempenho do sistema de otimização, incluindo a comparação de vários métodos de otimização. Para além disso, é realizada uma análise relativa ao melhoramento do fluxo de trabalho no estaleiro após a implementação do sistema, sendo enquadrada numa série de comparações entre as soluções originais (i.e., obtidas pelos métodos convencionais de dimensionamento) e as soluções otimizadas correspondentes. Em última análise, estes resultados ilustram o potencial e a importância da utilização deste tipo de tecnologias na gestão e otimização das terraplenagens.

CONTENTS

List of Figures

List of Tables

List of Symbols

A_0	Theoretical amplitude of a vibratory roller drum
C_i	Available rollers
C_{pl}	Cost associated with a production line
C_s	Cost associated with a single piece of equipment
$dcost$	Direct cost
e	Actual compacted layer thickness
F_C	Completed work fronts
F_i	Compaction fronts
F_R	Remaining work fronts
g_i	Gene indication the position of a roller in the embankment fronts
$icost$	Indirect cost
k	Efficiency factor
L	Roller drum width
M_0	Mass of the vibrating element excited by the eccentric mass
M_1	Total mass acting on the full width of a vibrating roller drum
m_e	Eccentric moment
Q	Nominal productivity
Q/L	Nominal compaction rate per drum meter width
Q/S	Compaction parameter for roller productivity estimation
Q_p	Actual/real/on-the-job productivity
R^2	Correlation coefficient
T_{pl}	Duration of work for a production line
T_s	Execution duration
V	Maximum roller speed for vibratory rollers or average speed for other roller types
V_c	Total volume of material associated with embankment front
V_m	Volume of material in current work front

“There is nowhere I cannot go,
yet I choose to be here.”
- *A Journey to the West*

Chapter 1

INTRODUCTION

1.1. MOTIVATION

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 (Figure 1.1). 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. This means that, during design phase of an earthworks project, balance must be found regarding excavation and embankment volumes.

Earthmoving is achieved by relying 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 (Gomes Correia and Magnan 2012; Gomes Correia et al 2014). The high number and type of heavy machinery required to complete earthwork tasks, allied to the nature of the tasks themselves, are two of the main reasons that make earthworks often incur the highest percentage costs and durations in road and railway construction projects. Effectively, these facts imply an increased importance regarding their optimization (Miao et al., 2011).

In this context, competitiveness in earthwork construction is mostly based on the achievement of lower execution costs and durations, currently more than ever considering the decreasing profit margins that companies and contractors are forced to undertake. Yet, these are inherently conflicting objectives, whose optimization implies the establishment of a trade-off by planners. In most cases, planners either find a possible solution based on their own experience, or settle for a random trade-off.

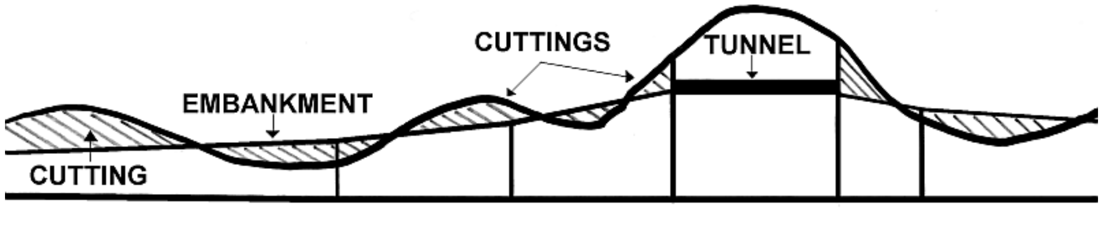


Figure 1.1 Schema of an earthworks project, including excavation (cuttings) and embankment areas

However, since recent advances in practices and technology enhance data collection in Civil Engineering construction, originating large databases of construction records, these can be used to fulfil the referred optimization objectives. This data can be associated with the construction layout, mass-haul diagrams, cut-fill routes and subsequent volumes of material involved in excavation and embankment construction, and even information regarding material compaction level or bearing capacity (White et al 2007; Rinehart and Mooney 2008). Soft computing (SC) techniques, namely data mining (i.e., neural networks) and metaheuristics (i.e., evolutionary computation), come forward as a way to use this data to optimize available resources in earthworks, being versatile enough to adjust to different environments and dynamic site conditions. Aiming at a maximization productivity (thus minimizing execution times) and minimization of costs, while ensuring the completion of the work within time and cost estimates, effective planning in these constructions is essential. In this point of view, both the allocation of available resources and the selection of the best equipment fleet for the work at hand are major factors to achieve these objectives.

There has been reasonable development regarding the optimization of earthwork constructions, mainly in the form of equipment and operation modelling systems, in order to simulate site conditions and work sequence. Among these, most focus on planning and optimization during a project design phase (AbouRizk and Hajjar 1998; Marzouk and Moselhi 2002a; Marzouk and Moselhi 2002b; Cheng et al 2005; Zhang 2008; Cheng et al 2010), whereas few look to optimize the earthwork tasks themselves throughout construction phase (Moselhi and Alshibani 2007; Moselhi and Alshibani 2009). These types of systems are generally based on metaheuristics, using SC techniques such as evolutionary computation (i.e., genetic algorithms – GA) (Michalewicz and Fogel 2000), thus being considered intelligent earthwork optimization systems. Nevertheless, during design phase of earthwork construction, the information regarding key factors with a direct influence on cost and duration (i.e., equipment productivity) of the planned tasks is frequently very scarce or even inexistent. Although a limited number of these systems have the ability to calculate features, such as the real equipment productivity during the construction work itself, the lack of accuracy of this information during design phase can gravely hinder its ability to carry out accurate time/cost predictions, potentially leading to considerably losses. As such, bearing in mind the availability of past construction data, it becomes possible to resort to other type of SC techniques, such as those based on data mining (i.e., neural networks), to predict those factors during planning in design phase. These algorithms can, for instance, be used to adjust models which "learn" from past data and can then be used to predict how a particular set of features will behave in a similar or future situation. In this sense, research has originated some applications (Michalewicz et al 2007; Hola and Schabowicz 2010) which can be considered relevant for earthwork optimization, although these do not have the capability of optimizing an earthwork system by themselves. Nonetheless, the possibility of integrating both SC techniques (metaheuristics and data mining) has obvious advantages, especially considering how data mining can compensate for the design phase limitations of optimization systems by enhancing the predictive potential of the system. Finally, as the construction phase begins, it becomes possible to complement this information with the updated values of the estimated features, either by human observation of the construction conditions, or by using spatial technologies, such as geographic information systems (GIS) or global positioning system (GPS). This enhances the system with the flexibility to re-optimize itself in order to tackle the dynamic nature inherent to earthwork constructions.

1.2. AIMS AND METHODOLOGY

According to Hevner et al (2004), in the context of Information Systems, design science research should produce a “viable artifact in the form of a construct, a model, a method, or an instantiation”, to solve relevant problems in a given area, while consisting of a contribution to the academic world. Furthermore, this output should be attained by means of rigorous scientific methods during its construction, using the available means to reach the desired goals, and it should be subject to effective evaluation methods to validate its capacities. As such, in order to achieve the proposed objectives, it is essential to obtain a deep understanding of the necessary processes related with planning, monitoring and performing construction operations, namely concerning the general work flow and equipment capabilities for constructions involving earthwork operations. Moreover, the necessary methods, tools and technologies normally associated with not only the planning phase of this kind of construction, but also with the monitoring of construction sites need to be given full consideration and careful study.

This project is intended to create a knowledge base regarding the earthwork tasks, as well as an intelligent system, with capabilities to support the user’s decision-making. This system is based in modern optimization techniques (e.g. Evolutionary Computation), in order to obtain an adequate modelling of real world problems, i.e. including the definition of features such as: multiple objectives, restrictions and dynamic environments, among others. The developed system is expected to be capable of performing both planning and monitoring constructions involving earthwork processes, while attempting to automatically retrieve equipment data, rates and productivities (namely for the compaction process). Therefore, developments associated with the envisioned system are related to:

- Given a certain number of available equipments, it should be able to select the best equipment fleet that can optimize the earthwork tasks, simultaneously minimizing costs and construction time;
- Given a predetermined deadline, it should feature the capability to select the necessary equipment fleet that allows for the completion of a project, or a set of tasks, in the available time;
- The capability to determine the optimum number of each type of equipment (optimal equipment allocation) in each work front; and
- Finally, since nowadays there is an environmental concern of using mainly landfill material derived exclusively from the construction site, it is essential to take into consideration the optimization of trajectories and transport distances between the material source sites and the work fronts, achieving a green construction, as well as proper management of the materials themselves.

Thus, the research strategy begins by means of an extensive and thorough literature study on earthwork construction. The developmental process entails adding new knowledge to previous experience on the field, allowing to better examine and understand the potentially relevant tools and technologies currently available. Moreover, the preliminary study allows for the assimilation of the conceptual features needed to build the earthwork optimization system. In summary, the steps for the completion of this project should include:

- In-depth research on earthwork optimization, identifying the technologies that can be used to achieve the above-mentioned objectives;
- Analysis of relevant systems comprising past attempts at solving the earthwork optimization problem, with focus on the used technologies, as well as on their subsequent strengths and limitations;
- Development of a novel earthwork optimization system, aiming to innovate the current state of knowledge, while surpassing the limitations of the previously analysed systems; and
- Application of the developed system for a real case study.

Considering the multi-disciplinary nature of the project, the end result is dependent on the successful combination of the technologies under study. However, the implementation of these technologies implies the integration of several tools as a basis for the system:

- By means of tools such as *R* (R Development Core Team 2011) and *Rminer* (Cortez 2010a), technologies related to data mining can used to create prediction models based on earthworks databases, such as the GTR compaction tables (SETRA and LCPC 2000), enhancing the system's planning phase capabilities;
- Using GIS software, such as ArcGIS (ESRI 2011), makes it possible to endow the system with the graphical and network analysis capabilities to support the user during the planning phase of a project; and
- Finally, through the use of optimization tools such as *R* or *Matlab* (The MathWorks Inc. 2014), and by integrating them with the previously discussed technologies, the system becomes able not only to identify and select the best found solutions for the near-optimal equipment fleet to be used in the planning phase, but also to re-optimize the complete system during construction phase in order to deal with unknown or unforeseen circumstances regarding the equipment fleet and work flow in the construction site. Thus, technologies such as Modern Optimization can be used as the centre of both the planning and monitoring features of the system.

By successfully implementing all these different technologies, the expected result will originate an innovative system aimed at supporting the decision-making process associated with earthwork processes in Civil Engineering environments.

1.3. OUTLINE OF THE THESIS

The thesis is divided into the following six Chapters, excluding Introduction (Chapter 1), which are consistent with the steps outlined in the previous subsection:

- Chapter 2 provides a background on earthworks construction. The allocation of the mechanical equipment associated with earthworks is defined as comprising an optimization problem, to which conventional design does not have a definitive solution. This Chapter also includes the methodologies and concerns related to conventional earthworks design, as these represent a fundamental base of knowledge for the understanding of the remaining thesis.
- Chapter 3 presents the study of the relevant technologies, which have the potential to deal with complex and dynamic environments, as is the case of earthworks construction. Subsequently, an analysis on the available tools that can potentially allow for the integration of the studied technologies is also carried out.
- Chapter 4 includes a study regarding previous attempts on using the mentioned techniques. The focus is on the concretization of the previously developed systems, emphasizing the used technologies. In order to achieve this purpose, an analysis is carried out with the intention of sorting the systems into their typologies, once their characteristics and aims will allow a deeper understanding of the subsequent strengths and limitations of the resulting systems.
- Chapter 5 consists of the development of the novel intelligent earthwork optimization system. This includes its architecture, elaborated on the development of the individual modules that comprise the system. The Chapter presents the new system as a means of filling the gap left by previous systems' limitations, particularly respecting the attempts at handling the earthwork optimization problem.
- Chapter 6 is related with a twofold application of the system for a case study. The two aspects that are addressed originate from the multi-disciplinary nature of this project. As such, application must be carried out bearing in mind the performance of the system, framed into the Information Systems area, as well as the quality of its output, which is more directly related to the Civil Engineering area. Moreover, since the required data is available, a comparison of the results obtained by the system and those obtained by conventional design is also carried out.
- Finally, Chapter 7 refers to the conclusions that were drawn during and after the development of the system. A description of the system's capability to improve the design and construction is provided. The discussion of results includes the potential of the system, acknowledging the tool's capability to support and improve earthworks design in the point of view of a company, and illustrating how its usage can have a significant impact on an economic perspective. This Chapter also draws some directions for future research and development on the subject under study.

Chapter 2

EARTHWORKS

2.1. INTRODUCTION

This Chapter intends to highlight the definition of earthworks in a Civil Engineering point of view, as well as to describe in detail the workflow and tasks that comprise it. In an optimization context, the duration and cost of earthwork tasks is optimal if the usage of resources (i.e. mechanical equipment) is ideal and site conditions allow it. In this point of view, knowledge of factors such as productivity of equipment versus the cost of their usage is paramount to allow for proper optimization. In turn, productivity is strongly influenced by the type of material that is used during construction, as well as by aspects related to site conditions, equipment specifications or operator skill. As such, Section 2.2.1 attempts to examine and detail the main elements that have an impact on the productivity of equipment for the given conditions, along with the methodologies used in conventional design to determine this productivity. Furthermore, since the optimization of earthworks is not limited to economic and technical aspects such as project duration and costs, insights into some current environmental concerns, such as carbon emissions in earthworks, are also provided, as they will be of interest to the remaining Chapters.

2.2. EARTHWORKS WORKFLOW

Due to constructional and environmental concerns in linear constructions (e.g. road, railway or airway constructions), designers generally attempt to limit the material usage to that which is available within the construction site. In other words, earthworks design is carried out in such a way that the totality of the material required for building each embankment in the construction site is originated in its excavation fronts, avoiding as much as possible the transportation of geomaterials from external sites (i.e. borrowing fronts) to be used in embankments. In practice, this means that a balance between excavation and embankment material volumes is sought during design phase. This results in an early definition of which areas will be excavated, so as to lower the actual topography to the target height, and where embankments are going to be built, in order to increase the height of the actual topography towards the design height.

As such, bearing in mind that the excavation and embankment areas are predefined, the earthworks process is generally regarded as a sum of sequential tasks which flow from excavation towards embankment tasks. Since the goal is to alter the surface/topography of a target area, the tasks generally associated with earthworks mostly follow the following sequence, excluding supplementary tasks, such as treatment of geomaterials (e.g. using cement and lime):

- **Excavation** of geomaterials. Resulting geomaterials are loaded onto trucks for transportation to embankment fronts.
- **Transportation** of excavated geomaterials from excavation fronts to embankment fronts. Materials are unloaded at a specific location inside an embankment front so as to allow for spreading.

- **Spreading** of previously unloaded geomaterials into layers to allow for compaction. Some geomaterials (e.g., soil-rockfill mixtures) require specific techniques to produce a properly spread layer.
- **Compaction** of layers of previously spread geomaterial. This task is generally the most complex of the process, requiring quality control procedures to be carried out during and after its execution.

These sequential tasks are achieved by relying on heavy mechanical equipment. The equipment varies depending on the task, and combinations of different equipment types are often used for each task. Table 2.1 summarizes and exemplifies the equipment types used in general, associating them with the respective task and other tasks they may have direct influence on. Intermediate tasks, such as material treatment or layer wettening during spreading and compaction, are not explicitly included in this simplified representation of earthworks, as they are performed according to material requirements and often depend on external factors (e.g. atmospheric conditions). In the particular case of layer wettening, this task is considered to be included in the spreading task throughout this work, specifically for productivity calculation purposes in Chapters 5 and 6.

Table 2.1 Earthwork tasks and equipment

Task	***Equipment***	***Interactions with other equipment***	***Type of interaction***
Excavation	Excavators, loaders, backhoe loaders	Transportation equipment	Loading
Transportation	Dumper trucks, articulated trucks	Excavation and spreading equipment	Loading (excavation fronts) and unloading (embankment fronts)
Spreading	Bulldozers, motor graders	Transportation and compaction equipment	Spreading unloaded geomaterial into layers
Compaction	Rollers (e.g., vibratory and tire rollers)	Spreading equipment	Compacting geomaterial layers

Since the mechanical equipment is generally limited in this type of construction, conventional design is usually focused on grouping a number of each type of equipment into predefined plants, depending on the available equipment and site conditions (e.g., material types, space restrictions, atmospheric conditions). For instance, a standard plant may consist of one excavator in an excavation front, supported by four dumper trucks, which transport the material to an embankment front with one bulldozer and one roller (Figure 2.1). Depending on the equipment plant (i.e., available equipment in a construction site) and site conditions, multiple plants of this kind can be working simultaneously in different excavations and embankment fronts. Furthermore, the amount of equipment allocated for each task may be adjusted so as to better adapt to the specific case of each front. Under conventional theories, adding or removing equipment to a specific task will linearly increase or decrease its productivity, respectively. For example, this may consist of adding additional trucks to compensate for a long hauling distance between an excavation and an embankment front; or allocating extra excavators to a front in order to increase its output, allowing the support of more than one embankment front. As work progresses, the available geomaterial in excavation fronts or the required geomaterial volumes in embankment fronts will gradually become exhausted or completed, respectively. When this happens, the equipment associated with the completed fronts is subsequently reallocated into another front and adjusted to its requirements and restrictions (e.g., increasing or decreasing amount of allocated equipment). This is carried out successively until all fronts are at the target height and the earthworks project is completed.

In this context, the time that each plant requires to complete their tasks is mostly dependant on their actual (i.e., also referred to as real or on-the-job) productivity, which, albeit being determined individually for each piece of equipment and its working conditions (as discussed further ahead on this Chapter), is strongly influenced by the productivity of the rest of the allocated plant. Bearing in mind the previous example, a truck plant is unable to transport more geomaterial volume than that which is excavated by their associated excavation plant, even if its nominal productivity rate is higher than the productivity of the excavators. Conversely, spreading equipment cannot work at its maximum potential in terms of productivity if the flow of geomaterial transported to the embankment front is insufficient and so on. Taking this into account, the productivity of the equipment allocated to a task is always conditioned by the productivity of the equipment allocated to the previous tasks. While adding more equipment to a specific task may increase its productivity, its maximum work rate cannot exceed that of the task that precedes it.

The actual productivity, Q_p (m^3/h), of all equipment in an earthwork construction is usually considered in terms of its nominal productivity, Q, (i.e., potential work rate, determined in function of equipment specifications) corrected by an efficiency factor, k (Equation 2.1).

$$Q_p = Q \times k \tag{2.1}$$

The purpose of the latter is mainly to take into account aspects that may hinder the performance of a piece of equipment, such as operator efficiency, material characteristics, haul road conditions, among others. However, the efficiency factor of a piece of equipment or a plant will also diminish as a result of delays or insufficiencies related with low productivity in previous tasks. Since working at 100% efficiency cannot be achieved continuously even in ideal conditions, for which the nominal productivity is determined, efficiency factors usually vary between 0,3 – 0,8 for earthwork equipment. This is also referred to as a number of minutes that a piece of equipment will work per hour, e.g., an efficiency factor of 0,75 will often be referred to as a 45 minute work hour in earthwork guides (SETRA and LCPC 2000; Caterpillar Inc. 2011).

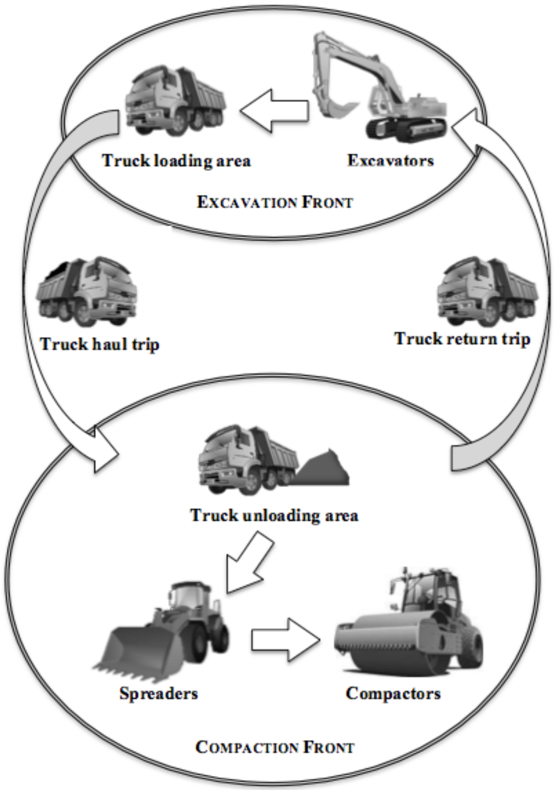


Figure 2.1 Earthworks equipment workflow

2.2.1. Equipment productivity

Excavation

In general, the factors that affect the productivity of most equipment in earthworks are related to equipment specifications, site and material conditions and operator skill. Depending on the equipment type, additional factors might be relevant in assessing the productivity rates. Naturally, the relevant specifications are different for each equipment type, as are the site and material conditions with impact on the productivity of a plant. As far as the excavation task is concerned, Figure 2.2 represents a diagram that depicts the most relevant factors influencing the real work rate of excavators in earthworks. Whereas equipment specifications, such as operating weight and bucket capacity, as well as some geomaterial characteristics, such as type and moisture state (i.e., water content), are the main aspects that determine the nominal equipment work rate, aspects like the operator skill, space restrictions and atmospheric conditions directly affect the efficiency at which an equipment can carry out its work.

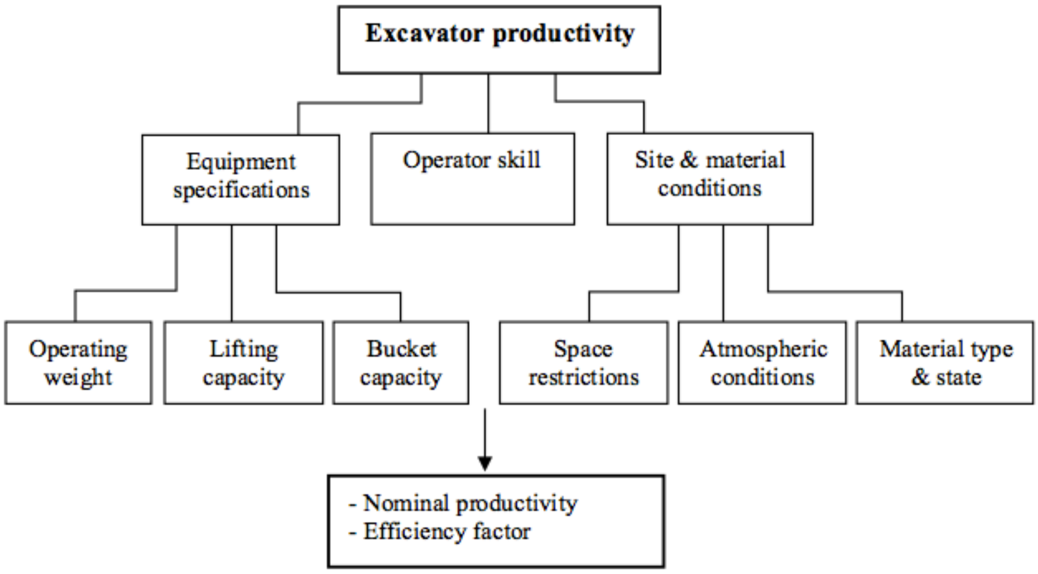


Figure 2.2 Relevant factors affecting excavator productivity in earthworks

More specifically, the nominal productivity of an excavator depends on its maximum bucket load and cycle time. In turn, the former is directly related to the equipment lift capacity and its operating weight, while the latter is linked to the potential positioning of excavators in the excavation front (e.g., excavator height in comparison with the height of the material being excavated, as well as with the angle/distance to the loading area/truck), operator skill and atmospheric conditions.

The choice of maximum bucket size for an excavator is limited by its lift capacity. An excavator's lift capacity depends on its weight, centre of gravity, lift point position and hydraulic capability. Furthermore, for any given lift position, the excavator lift capacity is limited by its hydraulic capacity, which changes depending on the equipped boom, stick and bucket, or tipping stability. According to design specifications (Caterpillar Inc. 2011), an excavator is considered to be at the tipping point when the bucket weight acting at the centre of gravity causes the rear rollers to lift clear of the track rails. Thus, the tipping load is defined as the load producing a tipping condition at a specified radius. The load radius is measured as the horizontal distance from the axis of upper structure rotation (before loading) to the centre of vertical load line with load applied (Figure 2.3, dimension A). The rating height is based on the vertical distance of the bucket lift point to the ground (Figure 2.3, dimension B).

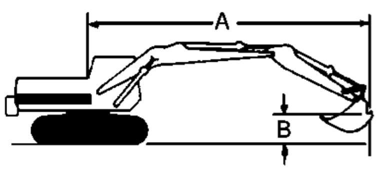


Figure 2.3 A. Load radius; B. Rating height

The rated load is established using the load radius and the rating height. Ratings for the ability of a specific machine attachment to lift a load slung from the designated bucket are defined as follows:

- The rated load must not exceed 75% of the tipping load;
- The rated load must not exceed 87% of the excavator hydraulic capacity. This means the machine should be able to lift 115% of the rated load; and
- The rated load must not exceed the machine's structural capability.

The rated load is directly related to the volume of material in the bucket in each loading cycle (i.e., bucket payload). Buckets are usually rated in terms of struck and heaped capacities. The struck capacity corresponds to the volume that is enclosed within the bucket geometric boundaries, disregarding any material above the strike off plane. The heaped load corresponds to the volume of the struck load plus the volume of heaped material above the strike off plane (Figure 2.4), considering an angle of repose of 1:1.

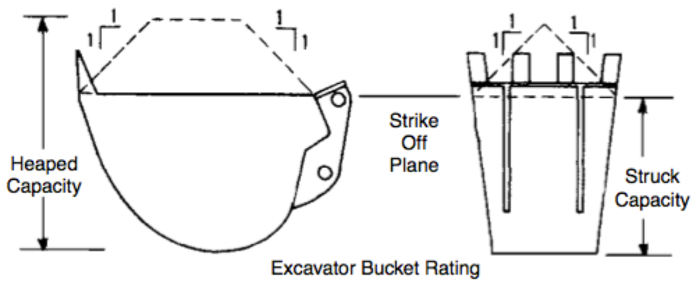


Figure 2.4 Bucket struck and heaped capacities

However, an excavator bucket payload is not only dependent on bucket geometric characteristics, but also on curl force and soil characteristics. The latter are usually taken into consideration by correcting the heaped bucket capacity with a fill factor (Equation 2.2). Fill factors for several types of materials are listed in Table 2.2.

$$Average\ Bucket\ Payload = Heaped\ Bucket\ Capacity \times Bucket\ Fill\ Factor \qquad (2.2)$$

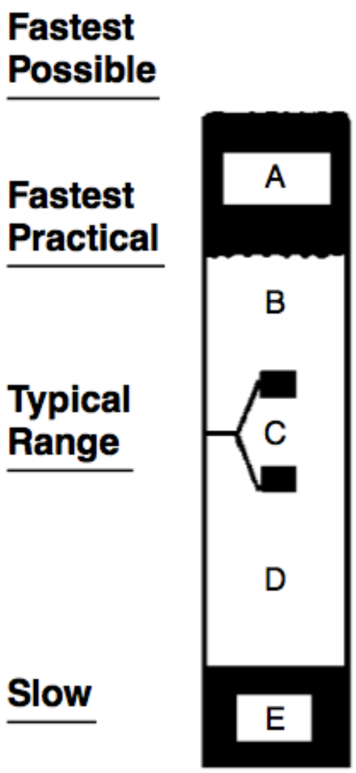
Table 2.2 Fill factors in function of material type

Material	***Fill factor range***
Moist loam or sandy clay	100-110%
Sand and gravel	95-110%
Hard, tough clay	80-90%
Rock – Well blasted	60-75%
Rock – Poorly blasted	40-50%

As mentioned, the equipment cycle time for given site conditions is also a critical aspect of excavation productivity. The digging cycle of an excavator is composed of four segments:

- Load bucket;
- Swing loaded bucket;
- Empty bucket; and
- Swing empty bucket.

The total excavator cycle time depends on equipment dimensions, since smaller excavators can cycle faster than larger ones, and site conditions. The more severe site conditions are, the longer the cycle times become. For instance, as materials become harder to dig, the difficulty of filling the bucket increases, and as the excavation point becomes deeper the bucket has to travel farther, increasing cycle time. The positioning of the unloading point (e.g., truck or spoil pile) in comparison with the excavation point in terms of turning angle also has a strong influence on the cycle time. If a truck is directly beside the excavation point and in a lower ground, very fast cycles are attainable, where otherwise if a truck is located 180º and above the excavator, the opposite will occur. Figure 2.5 summarizes how cycle times can vary depending on site conditions, giving practical examples.



A. Easy digging (unpacked earth, sand gravel, ditch cleaning, etc.). Digging to less than 40% of machine's maximum depth capability. Swing angle less than 30°. Dump onto spoil pile or truck in excavation. No obstructions. Good operator.

B. Medium digging (packed earth, tough dry clay, soil with less than 25% rock content). Depth to 50% of machine's maximum capability. Swing angle to 60°. Large dump target. Few obstructions.

C. Medium to hard digging (hard packed soil with up to 50% rock content). Depth to 70% of machine's maximum capability. Swing angle to 90°. Loading trucks with truck spotted close to excavator.

D. Hard digging (shot rock or tough soil with up to 75% rock content). Depth to 90% of machine's maximum capability. Swing angle to 120°. Shored trench. Small dump target. Working over pipe crew.

E. Toughest digging (sandstone, caliche, shale, certain limestone, hard frost). Over 90% of machine's maximum depth capability. Swing over 120°. Loading bucket in man box. Dump into small target requiring maximum excavator reach. People and obstructions in the work area.

Figure 2.5 Cycle time vs. site conditions chart (adapted from Caterpillar Inc. 2011)

Finally, having determined the average bucket payload and the cycle time for each excavator, it becomes possible to determine its nominal productivity, which can then be corrected into its on-the-job work rate by correcting it with the efficiency factor. The nominal productivity, usually measured in m^3/h, is regarded as the average bucket payload (determined by Equation 2.2) multiplied by the number of cycles per hour, as shown in Equation 2.3, while the on-the-job work rate, can be determined by applying the associated efficiency factor, as mentioned in Equation 2.1.

$$Q = Cycles\ per\ hour \times Average\ Bucket\ Payload \tag{2.3}$$

Hauling

Resembling the excavator case, truck productivity also depends on the general mentioned aspects: equipment specifications, site conditions and operator skill. In particular, the hauling distance is traditionally regarded as one of the main factors that affects the duration of hauling cycles. Even though the productivity of a truck plant is not limited to that factor alone, its optimization must include the decision of the trajectory taken in the hauling cycles, reducing cycle time, fuel usage and carbon emissions. In this context, tools such as geographic information systems can be used in the determination of shortest routes between loading and unloading areas, as further explored in Chapter 5. Moreover, since a truck is loaded by excavators, there is another decisive aspect that determines the productivity of a truck plant that was not included in the previous task: the productivity of excavators (Figure 2.6).

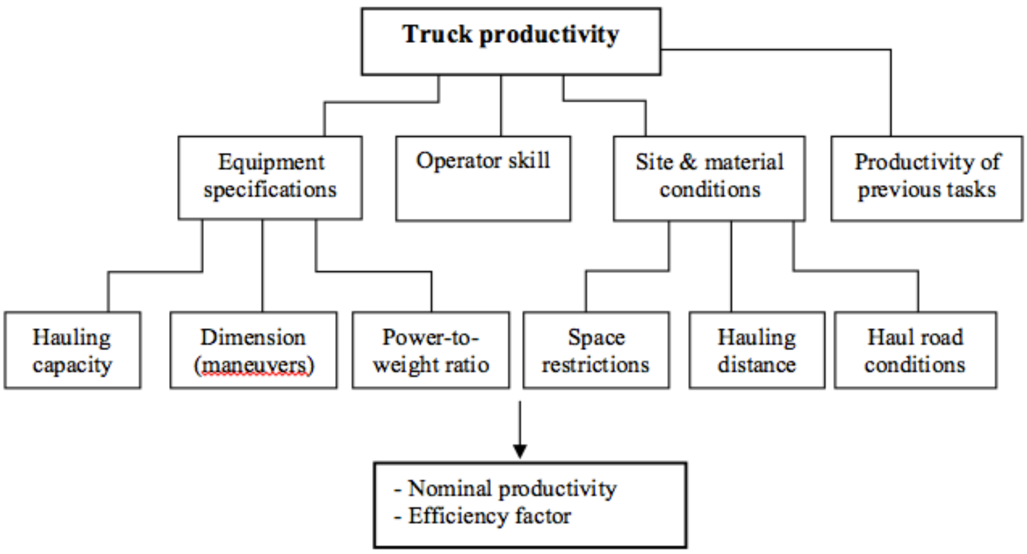


Figure 2.6 Relevant factors affecting hauling productivity in earthworks

In fact, if an excavation plant does not have a high enough work rate that it is able to load enough material onto trucks after each hauling cycle, the productivity of the hauling plant will be limited by the productivity of the excavator plant. This means that the work rates of both plants must be matched to each other, so as to avoid the unnecessary waste of resources and manpower (truck and excavator operators). For instance, if there are too many trucks for an excavation plant to load at each cycle, the former will have to incur in idle time while waiting for other trucks to be filled. Not only the waste of resources, but also of time, fuel, and manpower could be avoided by proper design. A mitigation measure to this problem could be either to increase the excavation plant's output during construction phase, if site conditions allow it and there is enough available equipment, or otherwise allocate some of the trucks to other fronts where they can be used at their full potential.

In order to match the productivities of both tasks, conventional design usually makes use of the concept of matching factor (Morgan and Peterson 1968), MF (Equation 2.4). The MF represents the ratio of actual truck arrival rate to loader service time. When the match factor equals 1, the operation is referred to as the ideal condition for determining the number of machines and the cycle time of equipment. If MF < 1, the operation indicates that less than the ideal number of hauling equipment is employed, while if MF > 1, it indicates that there is an exccess of haulers in comparison to the number the operation requires. Subsequently, the efficiency of a truck plant will stop increasing if MF > 1. However, this theory is limited to homogeneous truck fleets, which is not always possible, and the cycle times must not take into account any idle or waiting times.

$$MF = \frac{Number\ of\ haulers \times Loader\ cycle\ time}{Number\ of\ loaders \times Hauler\ cycle\ time} \tag{2.4}$$

The maximum nominal productivity of a truck depends on its maximum capacity (how much material it can carry per trip) and all the factors that affect the ammount of trips (i.e., cycles) it can do in a time window, for which one hour is usually taken as a reference (Equation 2.5). These include factors such as hauling distance, haul road conditions, excavator productivity, operator skill, space restrictions and equipment dimensions (e.g., for performing the necessary maneuvers both in loading and unloading areas). This nominal productivity is often corrected by the MF itself for cases in which MF < 1, as depicted in equation 2.6.

$$Q_{max} = Trips\ per\ hour \times Load\ volume\ per\ trip \tag{2.5}$$

$$Q = Q_{max} \times MF\ for\ MF < 1$$

$$Q = Q_{max}\ for\ MF \geq 1 \tag{2.6}$$

Finally, the real productivity of a truck plant can be determined by multiplying the nominal productivity by the efficiency factor, as mentioned earlier (Equation 2.1). However, it is important to note that the efficiency ratio for trucks is usually lower that its counterpart for other equipment types, often averaging between 0.4-0.6, and it can suffer significant variations even throghout a single construction project. This is due to the difficulty of controlling site and working conditions for each case. In general, even if a good design in terms of allocation throughout the construction phase is assumed, the efficiency ratio for trucks will still suffer strong variations with aspects such as:

- The ammount of necessary maneuvers in loading and unloading areas increases, which can be variable depending on the space restrictions and the ammount of equipment already allocated in each front.
- Hauling road conditions, including gradient, ground surface homogeneity and rolling resistance, which may hinder the maximum circulation speed for certain types of trucks.
- Operator skill, truck type and truck dimensions. Trucks may be classified according to several factors, such as type of engine, number of gears, number of wheels and axles, method of dumping (read-dump, side-dump, bottom-dump) and, obviously, truck capacity. Two trucks with similar dimensions may have different payloads due to different design. For instance, articulated dump trucks may have an easier time fitting into more turning points than rigid frame trucks.
- The percentage of the capacity of each truck that is actually loaded in each cycle, since the average payload of most trucks can vary depending on the loading conditions, such as available material, ease of excavation of each material type or ammount of trucks waiting in line. Furthermore, the determination of the ammount of material that is loaded into a truck must take into account the increase in void ratio that is a consequence of excavation, which often has effects on the density of the material. As a result, the actual payload of a truck quite often proves to be less than what is claimed by manufacturer specifications. Moreover, for some hauling routes, spilling during haul trip might be significant.

Spreading

Similarly to the truck case, one of the main factors with influence on the work rate of a spreading plant is the productivity of previous tasks. Obviously, this includes the productivity not only of the truck plant, but also of the excavation plant indirectly, since the work rate of the former depends on the work rate of the latter. Like other earthworks equipment, the remaining factors that generally affect productivity comprise the specifications of a bulldozer or tractor, the site and material conditions and the operator skill, as depicted in Figure 2.7.

Since the outcome of a bulldozer is measured in how much geomaterial it can spread per hour, the dimensions of the dozing blade and the traction power capabilities of the tractor are essential to determine its nominal productivity. In this point of view, properly matching tractor and dozer is a basic requirement for maximizing production.

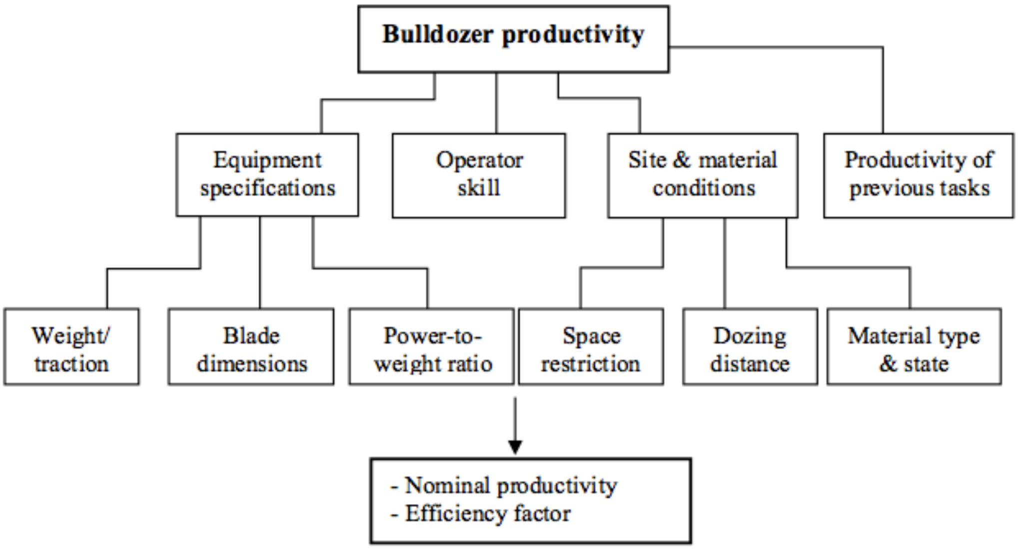


Figure 2.7 Relevant factors affecting bulldozer/tractor productivity in earthworks

In order to achieve this objective, one must evaluate the limitations of the available tractors as well as the materials to be moved. Regarding the former, the weight and horsepower of the equipment represent the main factors that determine their ability to push. However, site conditions, such as terrain and underfoot conditions, may limit the ability of a tractor to use its weight and horsepower. In terms of the material to be moved, dozer performance will vary characteristics such as:

- Particle Size & Shape – Larger particle sizes increase the difficulty for a cutting edge to penetrate the material. In addition, sharper edges also represent a natural resistance to the rolling action of a dozer blade, requiring higher horsepower to move.
- Void ratio – A low void ratio for a soil means the individual particles have more of their surface area in contact with other particles. Such a soil is generally heavy and will be hard to remove from the bank state.
- Water Content – In most materials the lack of moisture increases the bond between particles and makes the material difficult to remove from the bank state. High moisture content makes dozing difficult because the material is heavy and requires more force to move. Optimum moisture reduces dust and offers the best condition for dozing ease and operator comfort.

In view of this, the nominal productivity of bulldozers can be calculated based on numerical and empirical studies and it is often supplied by the manufacturers in tables or curves such as the ones presented in Figure 2.8.

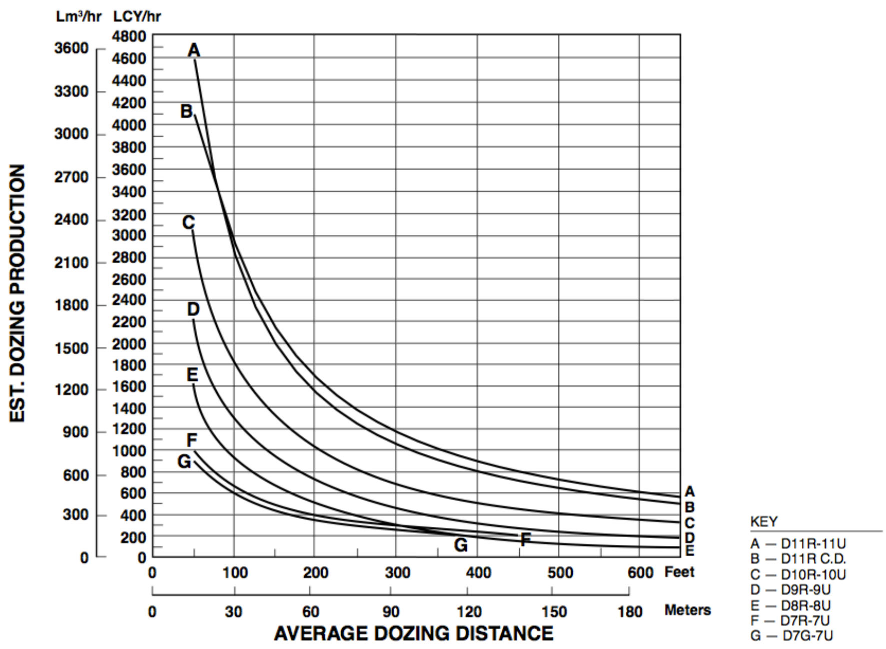


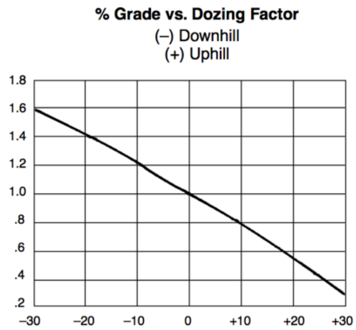
Figure 2.8 Example of bulldozer nominal productivity curves and associated models (Caterpillar Inc. 2011)

In the case of the presented productivity curves, the real productivity of bulldozers under specific conditions can be determined by correcting the nominal productivity by site condition factors, including the efficiency factor (exemplified in Table 2.3), as shown in Equation 2.4 (which falls under the same concept as Equation 2.1). These take into account the more obvious aspects, such as the material type and operator skill, and also other site conditions, such as visibility and terrain grade.

$$Q_p = Q \times Site\ condition\ factors \tag{2.4}$$

Table 2.3 Site condition factors for track-type bulldozers

OPERATOR	
Excellent	1.00
Average	0.75
Poor	0.60
MATERIAL	
Loose stockpile	1.20
Hard to cut; frozen	0.80
Dry non-cohesive material or very sticky material	0.80
Rock, ripped or blasted	0.60

TECHNIQUE	
Slot dozing	1.20
Side by side dozing	1.15-1.25
VISIBILITY	
Dust, rain, snow, fog or darkness	0.80
EFFICIENCY FACTOR	
50 min/h	0.83
40min/h	0.67
GRADES	

Compaction

Compaction is typically the most complex task of the earthworks process. In fact, not only does the interaction between equipment and geomaterial include a considerably higher amount of variables, but also the chosen methodologies and techniques can result in completely different work rates for the rollers. Figure 2.9 shows a summary of the most relevant factors affecting roller productivity.

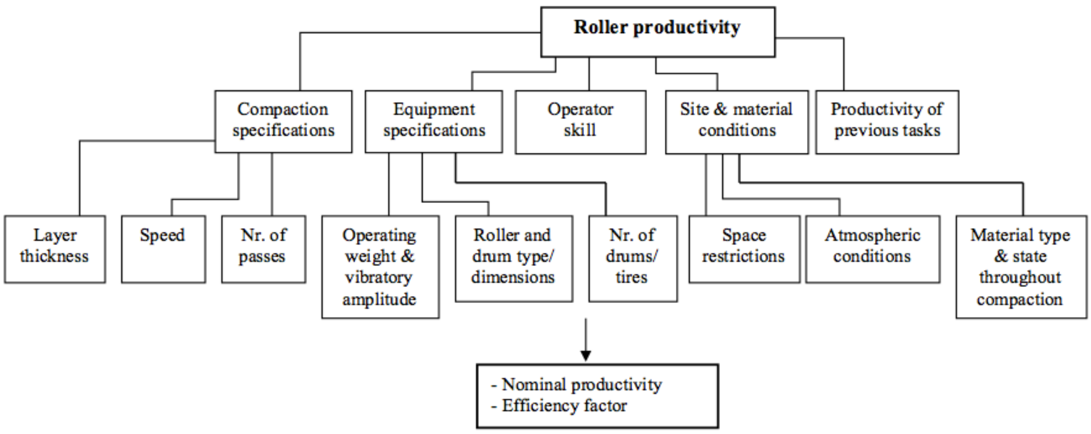


Figure 2.9 Relevant factors affecting roller productivity in earthworks

Thus, in order to accurately determine the productivity of a roller or a roller plant, it is essential to characterize both the materials (in terms of material type and conditions *in situ*) and the compaction equipment (taking the roller specifications as a reference) being used. Each roller will have different productivity rates associated with each geomaterial, as well as with the compaction specifications for the task. Several compaction guides have attempted to capture the complex variables associated with the compaction task (British Standards Institution 1981; Transportation Research Board and National Research Council 1990; SETRA and LCPC 2000). Since compaction productivity is even more influenced by materials and equipment specifications than any other task, these guides usually follow similar concepts and methodologies, attempting to quantify the difficulty of compacting each type of material commonly found in the associated country. This includes a detailed evaluation of geomaterial characteristics, both in terms of mechanical properties and plasticity, followed by a description of compaction tecniques and methodologies, which are different for each type of compaction equipment.

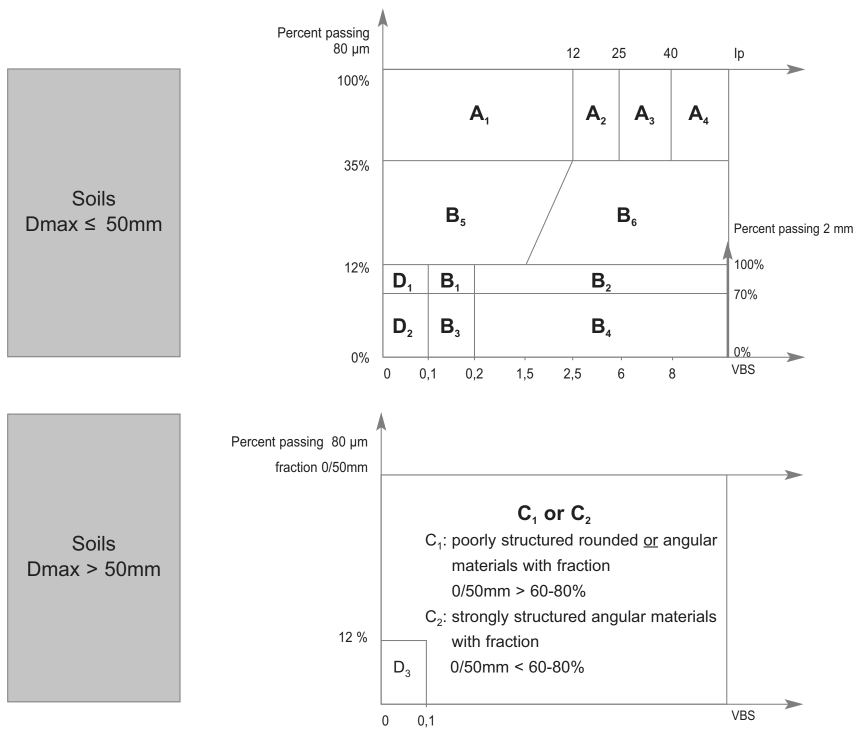
Among these, the French *Guide des Terrassements Routiers*, GTR, (SETRA and LCPC 2000) features a broad classification of geomaterials and rollers based on empirical data and has proven its effectiveness by being used in a wide range of countries and conditions. According to the above-mentioned guide, in order to properly estimate the productivity of a roller given a specific geomaterial, site conditions and compaction specifications, one should begin by characterizing the involved materials and available compaction equipment. This assessment must be detailed enough to allow for the determination of the necessary compaction energy to achieve good compaction results, depending on material type as well as condition *in situ*. Recalling **material classification**, the guide separates the classification parameters for both soils and rocks. The classification tables for soils and rocks are provided in Annex A and B, respectively. With respect to soils, three main topics are, in general, considered:

- Grain size characteristics, derived simply from the grain size analysis, is the initial parameter for soil classification;
- Clay characteristics, evaluated from the following three tests, which can then be used to specify the general group of the geomaterial:
  - Attenberg limits (plastic index);
  - Methyl blue absorption value of soil; and
  - Sand equivalent.
- In what concerns rock materials, the GTR classification divides the classes with relation to their geological labelling and the practical specifications of their use, ranging from R_1 – chalks – to R_6 – igneous and metamorphic rocks. The following two tests are worth emphasizing on account of their importance towards rock classification:
  - Fragmentation test; and
  - Degradability test.

Summarizing, when considering these general parameters, the GTR material classification can be considered by ranging from:

- A – Fine soils
- B – Sand and fine gravel (grit)
- C – Gravel soils with fines
- D – Soils and rock materials not influenced by water
- R – Rock materials
- F – Organic soils

These material classifications can still be separated in several sub-classifications, as shown in Figure 2.10, which briefly summarizes the classification methodology for geomaterials:



Rocks	Sedimentary rocks	Carbonate rocks	Chalk	R_1
			Limestone	R_2
		Argillaceous rocks	Marls, claystone, pelite, etc.	R_3
		Siliceous rocks	Sandstone, puddingstone, breccia, etc.	R_4
		Saline rocks	Rock salt, gypsum, etc.	R_5
	Igneous and metamorphic rocks	Granite, basalt, andesite, gneiss, schist, slate, etc.		R_6

Figure 2.10 Material Classification according to their nature (adapted from SETRA and LCPC 2000)

To complement material characterization the GTR also includes the state characteristics, especially considering the water content of the material. These also account for the environment in which the material will be used. The moisture state ranges from very moist to very dry. The normal state is the best condition for soil placement and compaction. Wet and very wet states are soils for which trafficability and compaction are difficult. Dry and very dry states are soils which are hard to compact and form stable embankment structures. This step results in the further classification of soil geomaterials according to their moisture state, and is made bearing in mind:

- The relation between its present water contents and the optimum water content from the Normal Proctor test;
- The soil consistency index, which represents the material's water content compared to the Attenberg limits;
- The measured IPI Index (consisting of the CBR Index measured with the soil's natural water content, and without surcharge);

Roller classification in GTR is restricted to rollers with a drum width of at least 1,30m. The classification tables for compaction equipment are compiled in Annex C. In general, rollers are divided into 5 categories:

- Pneumatic tyred rollers – classified depending on load per wheel. This type of roller is often ballasted up to twice their empty weight in order to obtain maximum wheel load recommended by the manufacturer;
- Smooth vibrating drum rollers and vibrating tamping rollers – Both are classified according to a parameter defined as $(M1/L)\sqrt{A0}$ and a minimum A0 value, where M1 is the total mass in Kg acting on the full width of the vibrating drum, L is the drum width in m, and A0 is the theoretical empty amplitude, calculated as $A0 = 1000\,(me/M0)$, in which me is the eccentric moment in mkg and M0 is the mass in Kg of the vibrating part excited by the the eccentric mass;
- Static tamping rollers – Static tamping rollers are classified according to the average static load per unit width of drum(s) with tamping feet;
- Vibrating plate compactors – Vibrating plate compactors are classified on the basis of static pressure under the plate Mg/S in kPa (Mg is the weight of the plate and S is the contact area between plate and soil).

Having carried out proper classification of existent geomaterials and available compaction equipment in a construction site, as well as moisture content and climatic conditions during work, it becomes possible to estimate the productivity of the compaction process. Being heavily dependent on the chosen compaction specifications, such as layer thickness, roller speed and number of passes, this information is presented in tables to allow for easier access. In these compaction tables, the following parameters are defined (SETRA and LCPC 2000):

- Q/S, a compaction parameter that is constant for all layer thicknesses, directly relating the layer thickness and number of necessary passes, Q/S = e/N;
- e, which represents the actual compacted thickness;

- V, standing for the maximum roller speed for vibratory rollers or average speed for other roller types;
- N, number of load applications (i.e., roller passes);
- Q/L, or nominal compaction rate per meter width, where L represents roller drum width, Q/L = 1000 x V x (Q/S). Real rate of compaction can be achieved by multiplying the nominal compaction rate by an efficiency ratio, k (usually between 0.5 and 0.75), which is connected to the number of hours per day in which the roller is actually performing compaction; and
- Applicable compaction code, which comes from soil use tables (based on moisture content and weather).

For most roller classes, the GTR shows, for a certain material, one single value for each of these parameters. However, in the specific case of vibratory rollers, the compaction tables present two columns, one regarding maximum compaction rate for a maximum velocity limited to 5 km/h, with a lower layer thickness; while the other concerns maximum layer thickness, with lower maximum roller speed (Figure 2.11).

Method		Class V3	
Code 2	Q/S	0.135	
	e	0.30	0.80
	V	5	2
	N	3	6
	Q/L	675	270

Figure 2.11 Example of the compaction table for B1 type soil and V3 class roller (adapted from SETRA and LCPC 2000)

According to this classification, should a nominal thickness for a specific job (e_{job}) with vibratory rollers fall between the two values in the compaction tables, it is possible to determine the optimum compaction conditions for that thickness. This should be done by calculating a new maximum speed for the roller, V = (V x e) / e_{job}, where the values V and e correspond to Vmin and emax (from the right hand column of the compaction table). The new speed will make it possible to determine a new rate per meter width Q/L = 1000 x V x (Q/S). The number of passes, N, is equal to e_{job} / (Q/S), thus making it possible to calculate an intermediate column for thickness values which fall between the ones presented in the tables. Having determined the nominal work rate of a roller for a given situation, making use of these compaction tables, it is then possible to correct it using the efficiency factor index (Equation 1), which takes into account aspects such as operator skill, space restrictions and productivity of previous tasks, resulting in the real work rate for each case. According to the aforesaid guide, in the case of compaction, the time that a roller takes to stop and accelerate back to compaction speed at the end of each pass should also be taken into consideration when selecting a value to the efficiency factor. Recommended design values for compaction equipment range from 0.50 to 0.75.

2.3. ENVIRONMENTAL CONCERNS IN EARTHWORK CONSTRUCTIONS

Traditionally, the main environmental concern of earthwork designers was directed at reutilization of materials for embankment construction. This includes not only the already referred usage of materials from the excavation fronts as much as possible, but also the treatment of these materials when their conditions are less than reasonable for use in embankments. As such, several soil treatment techniques have been developed in order to compensate for the drawbacks and limitations that might be implied in the employment of soils with specific characteristics. These techniques range from treatment with cement or lime, or both, in order to improve mechanical or plastic properties of the material (Silva et al 2013; Tinoco and Gomes Correia 2013).

In recent years, new environmental concerns regarding earthwork constructions have emerged. Such concerns are considerably wide, ranging from water economy to carbon dioxide emissions and waste control in construction phases, which go beyond the usual implemented environmental rules, such as the ecology and nature conservation or the regard for the landscape and townscape (London Health Comission and London Development Agency 2004). Most of these construction concerns have been taken into consideration in many recent major constructions, for instance during the construction and preparation for the UK Olympic Games, including (Olympic Delivery Authority 2007):

- Carbon: To minimise the carbon emissions associated with the construction of the Olympic Park and venues.
- Water: To optimise the opportunities for efficient water use, reuse and recycling.
- Waste: To optimise the reduction of waste through design, and to maximise the reuse and recycling of material arising during demolition, remediation and construction.
- Materials: To identify, source, and use environmentally and socially responsible materials.
- Biodiversity and ecology: To protect and enhance the biodiversity and ecology of the Lower Lea Valley, and other venue locations.
- Land, water, noise, air: To optimise positive and minimise adverse impacts on land, water, noise, and air quality.

It is easily inferred that carbon emissions and air quality are of critical relevance in earthworks tasks. Actually, as noted within the Environmental Statement, the key emission to air is the generation of dust from demolition, earthworks and construction activities.

Emissions from vehicles associated with construction sites can significantly add to levels of local air pollution, so it is important that best practices are adopted to reduce vehicle emissions. As such, several mitigation measures can be taken in order to minimize air quality impact (Greater London Authority and London Councils 2006):

- Developers can specify tax-exempt ‘red’ diesel with a sulphur content equivalent to ultra low sulphur diesel. This measure will automatically reduce particulate emissions by 30 per cent. Furthermore, fitting suitable after-treatment devices can reduce the remaining particles by at least 85 per cent. Also, considering that fine particles are of great concern to health, this is, therefore, a very effective way of reducing any health impacts to workers and sensitive receptors.
- In addition to local air pollutants, carbon dioxide is also emitted from vehicle exhausts. As it is a key gas linked to climate change, controls should be put in place to limit emissions; these controls will also help developers reduce fuel costs.
- No vehicles or plant should be left idle unnecessarily. Should any emissions of dark smoke occur (except during start up) then the relevant machinery should be stopped immediately and any problem rectified before being used.
- Reduce the number of vehicle movements through better planning.
- Set an appropriate speed limit on haul routes.
- Avoid use of diesel or petrol powered generators by using mains electricity or battery powered equipment where possible and if safety concerns can be overcome.
- Encourage developers to use consolidation centres to manage site deliveries. This will help reduce time wasted on searching for materials and the number of vehicles entering the site, and will have both congestion and emission benefits.
- Where construction sites are located near to waterways or railways it may be feasible for construction materials to be delivered or removed from the site using water transportation means, rather than by road. The obvious benefit is that it will reduce the number of trips made by vehicles on local roads, therefore reducing local emissions and disturbance to sensitive receptors.

2.4. FINAL REMARKS

Earthworks are comprised of several tasks that are carried out in sequence so as to shape the surface of an area to fulfil a specific purpose, such as foundation for structures of infrastructures. The progress of an earthworks project is influenced by a high amount of variables, which differ according to each task. Furthermore, the tasks feature a high level of interdependence as a result of their sequential nature: one task is unable to achieve its goals within reasonable time and budget if the previous tasks are not appropriately assembled to sustain it.

In this context, the optimization of earthworks must deal with the elevated complexity and high amount of variables inherent to the associated tasks. However, being one of the processes with higher costs and durations in many constructions (e.g., road and railway construction), the optimization of every task is paramount. By optimizing the whole earhtworks process, one is minimizing not only costs and durations, but also its environmental impact, by reducing fuel usage and carbon emissions.

Considering that there is availability of technology with the potential to be used for optimization of such complex processes (Chapter 3), this work explores the possibility to develop a robust system that combines several of these technologies in order to perform the global optimization of the whole earthworks process (Chapter 5), bearing in mind the concerns that have been discussed in this Chapter.

Chapter 3

TECHNOLOGIES AND TOOLS

3.1. INTRODUCTION

Following the background on the topic of earthworks in the previous Chapter, including the delineation of the problem being addressed in this study, the present Chapter compiles the available technologies used to achieve the proposed goals. As previously mentioned, bearing in mind the high complexity of the outlined optimization problem, conventional linear optimization or extensive search methodologies are infeasible. As such, metaheuristics, such as genetic algorithms or swarm intelligence, come forward as an alternative for dealing with large search spaces in complex environments, within reasonable computational requirements.

However, a good optimization output is not only dependent on the used method, but also on the proper knowledge of construction-specific restrictions and productivity factors. Considering this information is often unknown during project design phases, reliable estimation methods are required. Leveraging on the increasing availability of construction databases that stem from the development of information technologies, data mining (DM) emerges as a potential tool for dealing with this issue. Both the metaheuristics and the DM techniques mentioned in this Chapter are cathegorized as soft computing (SC) techniques, as they resort to solutions that attempt to deal with imprecision, uncertainty, partial truth, and approximation to achieve practicability, robustness and low solution cost in terms of computational requirements. For this reason, the term SC is often used in this work.

Finally, the advantages of geographic information systems (GIS) in this area of study should not be overlooked. In fact, given their capabilities to represent or map large areas (e.g., a road or railway construction site) using cartographic and geographic data, GIS are ideal tools to optimize traffic flow for different conditions (e.g., transportation equipment hauling and return routes) in a georeferenced environment, which would otherwise be impractical using other optimization methods. Furthermore, GIS feature capabilities related to geoinformation analysis and manipulation, as well as visual support for output visualisation, enhances their potential when used on their own or in conjunction with the mentioned technologies.

Subsequently, this Chapter focuses on the description of the features within these technologies that can be geared towards a global optimization of the earthworks process. Table 3.1 summarizes and structures the references associated with each type of technology, which are then described throughout the ensuing sections.

3.2. KNOWLEDGE DISCOVERY IN DATABASES

Knowledge discovery in databases (KDD) is an artificial intelligence branch focused on developing machines that display intelligent behavior, similarly to humans. One of the features that separates a human from a machine up to this point is the ability to learn from experience. KDD attempts to fill this gap, being defined as a nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data (Fayyad et al 1996a). The term nontrivial means that KDD process is not a straightforward procedure, instead consisting of an interactive and iterative process, involving numerous steps, in which the user must make fulcral decisions.

Table 3.1 Technology literature map

Technologies		
Data Mining	*Metaheuristics*	*Geographic Information Systems*
Brown and Kros 2003	Dorigo and Gambardella 1997	Denègre and Salgé 1996
Chang et al 2000	Glover 1986	Gutiérrez Puebla and Gould 1994
Chapman et al 2000	Glover and Laguna 1997	Heywood et al 2011
Fayyad et al 1996a	Haeser and Gomes-Ruggiero 2008	Martin 1991
Fayyad et al 1996b	Holland 1975	Peuquet and Marble 1990
Groth 2000	Kennedy and Eberhart 1995	Rowley and Gilbert 1989
Hastie et al 2009	Kirkpatrick et al 1983	Zeiler 1999
Hebb 1949	Michalewicz 1999	
Kim and Street 2004	Nahar et al 1986	
Li et al 2004	Petri 1966	
Liao et al 2012	Sujitjorn et al 2006	
Matsatsinis and Siskos 2002	Zadeh 1965	
McCulloch and Pitts 1943		
Musaev 2004		
Pearson 1908		
Rosenblatt and Cornell Aeronautical Laboratory 1958		
Santos and Azevedo 2006		
Turban et al 2004		
Vapnik et al 1997		
Vapnik 1998		
Walker 2007		

In this context, DM is formally considered part of the larger process, KDD, which represents the overall process of discovering useful knowledge from data (Fayyad et al 1996b). However, through the years the term data mining became more popular, being often used as a synonym of KDD. The latter is a multi-step process, including selection, pre-processing and processing of data, application of DM algorithms, interpretation and processing of knowledge. The rapid development of these methodologies can be traced to the increasing emergence of electronic data management methods, having successfully been applied to several different areas, such as marketing (Matsatsinis and Siskos 2002; Kim and Street 2004), manufacturing and production (Chang et al 2000; Musaev 2004) or health care (Li et al 2004; Walker 2007), as well as numerous applications in other areas (Turban et al 2004; Liao et al 2012).

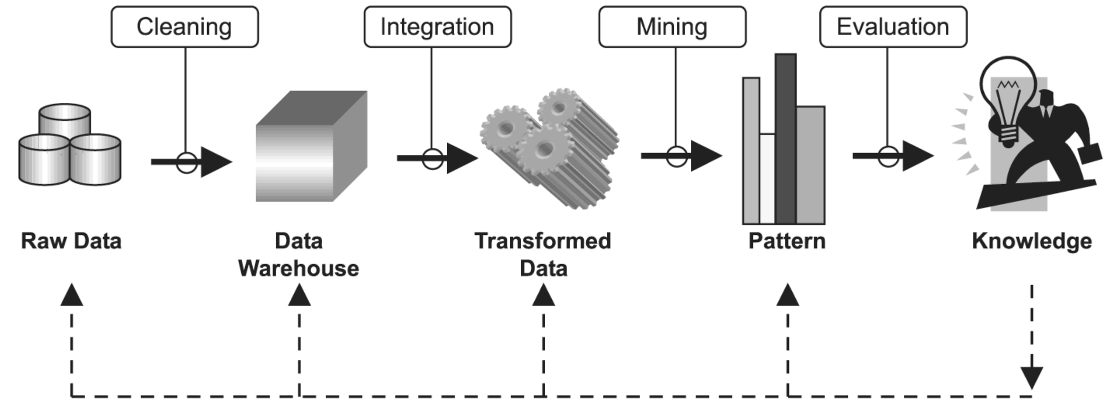


Figure 3.1 KDD process overview (adapted from Fayyad et al 1996b)

The first step of the process should consist of an understanding of the field associated with the problem and the definition of goals for the study, as well as the compilation of the available data on the subject. The ensuing database should then undergo a cleaning and preparation process, in which a subset of the original data is selected including exclusively relevant attributes. This step often requires a multidisciplinary team of experts, who can provide the necessary know-how to support the selection of relevant attributes. The cleaning process comprises the removal of noise and outliers, aiming to improve data quality (Brown and Kros 2003).

The integration, or transformation step covers the adaptation/modification of the data structure and type so as to allow for the application of the different DM algorithms, according to the requirements of the latter. As an example, a normalization of inputs and outputs to a zero mean with a standard deviation of one can be advantageous for the application of some DM algorithms.

The fulfillment of the previous stages allows for the application of DM algorithms to the applied data, corresponding to the mining step. This specific task is further discussed in subsection 3.2.1. Finally, the resulting patterns are then analysed and interpreted, potentially resorting to support methods or software (e.g., visualization tools), in order to obtain useful knowledge that can be used in the area of study. For instance, this knowledge can be used in a decision support process or incorporated in intelligent systems, such as expert of knowledge-based systems. It is to be noted that not all unearthed patterns are useful, thus demanding the careful analysis of the user or field experts. Moreover, it may be necessary to fall back to any of the previous steps to correct options and errors, increasing the quality of the final results. This shows how KDD is an interative procedure, where the quality of the results is dependent of the interaction between the discussed sequential steps and the user.

3.2.1. Data mining

When framed in the context of a methodology the DM process becomes easier to understand, implement and analyse (Santos and Azevedo 2006). As one of the most widely used and complete methodologies, the CRISP-DM (Cross Industry Standard Process for Data Mining) was first conceived in 1999 (Chapman et al 2000) and has already acquired a high level of acceptance in data mining projects second surveys to multiple users. The methodology has the advantage of being neutral concerning the industry where it is applied and the tools used, so users should use it as a framework to conduct the entire data mining process regardless of the business and the tools. The steps of CRISP-DM are translated into a hierarchical process, with a life cycle that is developed in six phases: business understanding, data understanding, data preparation, modelling, evaluation and implementation (Figure 3.2).

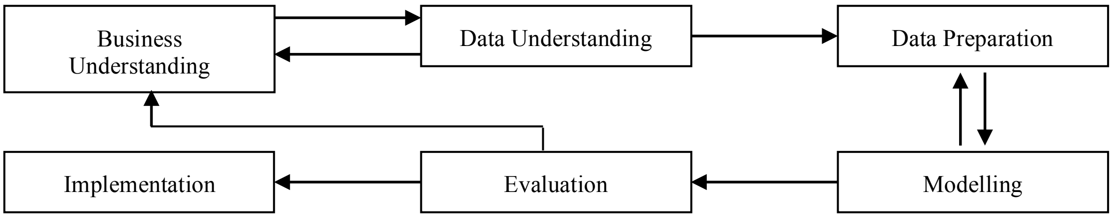


Figure 3.2 CRISP-DM methodology (adapted from Chapman et al 2000)

The description of the steps comprising CRISP-DM are as follows:

- **Business understanding:** identification and understanding of the project objectives and requirements from a business perspective. This is converted into a DM problem definition and a preliminary plan is proposed to achieve the goals;
- **Data understanding:** collection and analysis of the data in order to access its quality, discover first insights and detect subsets or trends. With this first data analysis, some hypotheses are formulated for hidden information;
- **Data preparation:** compilation of the final dataset that will be used during the learning phase (modelling) to build the DM model. Include the selection of the records and attributes from the initial raw data as well as its cleaning and transformation;
- **Modelling:** selection of the DM algorithms and optimization of its parameters in order to find patterns within the data;
- **Evaluation:** thorough assessment of all fitted models and revision of all previous steps in order to verify the achievement of business objectives; and
- **Implementation:** organization of the obtained knowledge and its implementations in order that it can be used by the user/customer.

DM tasks

There is currently a wide variety of DM algorithms and techniques, the choice depending mainly on the desired goal of the process. In general, the DM goals are either prediction or description (Figure 3.3).

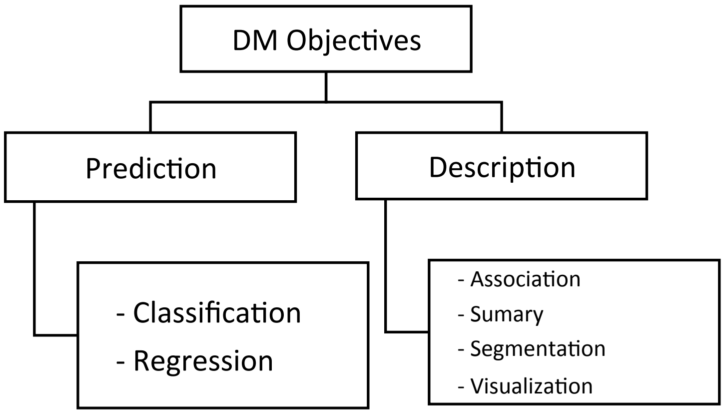


Figure 3.3 Data mining goals

Prediction aims to support decision-making by creating a model that can predict or estimate a value. The most common problems included in this category of forecasting are classification and regression. The first type of problem involves the discrimination of entries in a database into predefined categories (Fayyad et al 1996b). It consists of creating a set of models or functions that attempt to describe each class based on data analysis, so as to classify new entries according to a classification model. This is the most common objective of DM, usually carried out by techniques such as decision trees and artificial neural networks. As far as regression is concerned, its purpose is to predict unknown or future values of a dependent variable by creating a model from the known database (Fayyad et al 1996b; Hastie et al 2009). The most applied algorithms for this approach are multiple regressions, artificial neural networks and, more recently, support vector machines. In terms of applications, these are the extensive coverage, ranging from estimating the probability of a patient survival given the results of a series of diagnostic tests, to forecasting the market demand for a new product depending on the investment in its advertising.

Description finds its purpose in increasing the knowledge and understanding of the data, focusing on finding descriptive standards recognized or interpretable by humans (Santos and Azevedo 2006). The most common problems included in this goal are segmentation (also referred to as clustering in the bibliography), which aims to form groups of objects as homogeneous as possible with each other, for instance by seeking common features in the data that result in a potential clustering throughout various series; visualization, intended to describe complex information through diagrams, facilitating the visual representation and interpretation of patterns and trends; association, which identifies dependencies between variables, highlighting those that are considered more significant, widely used in transactional data analysis, such as marketing analysis; and summarization, aiming to find a representative compact description for a larger data set in order to allow for exploratory data analysis and automatic report generation.

Regarding the wide range of DM techniques, one of the conclusions that can be drawn is that there is no universal technique, since each one can be best suited to some problems than others, as suggested in the literature on the subject. Thus, the DM process is often referred to as an iterative process, in which several techniques are applied to the same problem in an attempt to identify which leads to better results (Fayyad et al 1996b). In this context, some DM techniques with more relevance for this work are discussed in the next subsection.

DM algorithms

Multiple Regression

Conceptually, a multiple regression (MR) (term first used by Pearson 1908) is a straightforward extension of the simple linear regression procedures. In a simple linear regression, a straight line (Equation 3.1) is adjusted to a set of data, attempting to model the relationship a scalar dependent variable, y, and one independent variable, x:

$$y = \beta x + \alpha \tag{3.1}$$

For multiple regression, more than one independent variable is included, being that for each new independent variable a new term is added to the model (Equation 3.2):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k \tag{3.2}$$

where β_i are coefficients for the independent variables to be adjusted, and the x_i are the values of the independent variables for the member of the population.

The process involves using data from a sample to obtain an overall expression of the relationship between the dependent variable, y, and the independent variables, x_i. This is done in such a manner that the impact of the relationship of the independent variables collectively on the value of y can be estimated. In other words, the value of a new dependent variable, ŷ, can be estimated using the coefficients β_i previously adjusted to a data set. Due to its additive nature, this model is easy to interpret and it is widely used in regression tasks.

Artificial Neural Networks

Artificial neural networks (ANN) are designed to mimic the human brain, being defined as "parallel distributed processors, consisting of simple processing units, which have a natural ability to store experimental knowledge and make it available for use" (Groth 2000). This definition implies that these are computational systems with a capacity to learn from their own use, allowing for the resolution of complex computational problems with nonlinear characteristics.

The development of ANN dates back to the early 1940s, designed by McCulloch, Pitts and Hebb (McCulloch and Pitts 1943; Hebb 1949). It experienced an upsurge in popularity in the late 1980s. This was a result of the discovery of new techniques and developments and general advances in computer hardware technology. Some ANN are models of biological neural networks and some are not, but historically, much of the inspiration for this field came from the desire to produce artificial systems capable of sophisticated, perhaps intelligent, computations similar to those that the human brain routinely performs, and thereby possibly to enhance our understanding of the human brain. Most ANN have some sort of training rule. In other words, they learn from examples (e.g., as children learn to recognize dogs from examples of dogs) and exhibit some capability for generalization beyond the training data. Neural computing must not be considered as a competitor to conventional computing. Rather, it should be seen as complementary, for the most successful neural solutions have been those which operate in conjunction with existing, traditional techniques (Table 3.2).

Table 3.2 Comparison between traditional computer techniques and ANN

Computers have to be explicitly programmed	ANN learn from examples
▪ Analyze the problem to be solved. ▪ Write the code in a programming language.	▪ No requirement of an explicit description of the problem. ▪ No need for a programmer. ▪ The neural computer adapts itself during a training period, based on examples of similar problems, even without a desired solution a specific problem. After sufficient training, the neural computer is able to relate the problem data to the solutions, inputs to outputs, and finally it is able to offer a viable solution to a brand new problem. ▪ Able to generalize or to handle incomplete data.

The brain is a collection of about 10 billion interconnected neurons. Each neuron is a cell that uses biochemical reactions to receive, process and transmit information. Each terminal button is connected to other neurons across a small gap called a synapse. A neuron's dendritic tree is connected to a thousand neighbouring neurons. When one of those neurons fire, a positive or negative charge is received by one of the dendrites. The strengths of all the received charges are added together through the processes of spatial and temporal summation (Figure 3.4a).

Neural computing requires a number of neurons, to be connected together into a neural network. Neurons are arranged in layers. Each neuron within the network is usually a simple processing unit which takes one or more inputs and produces an output. At each neuron, every input, p_i, has an associated weight, w_i, which modifies the strength of each input. The neuron simply adds together all the inputs and calculates an output, α, to be passed on (Figure 3.4b).

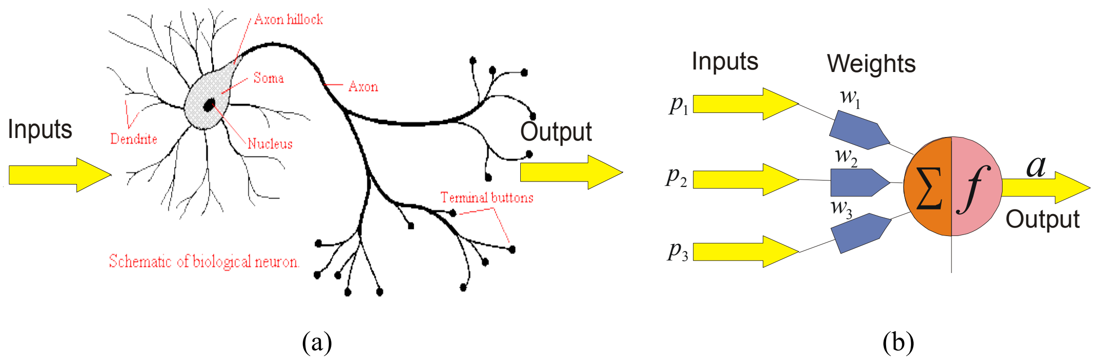


Figure 3.4 Schematics: (a) biological neuron; (b) neural network

The learning process of an ANN is based on specific algorithms with very well defined rules. In this context, there are two main methods, normally called paradigms, used for ANN learning process:

- **Supervised learning**: in supervised training, both the inputs and the outputs are provided. The network then processes the inputs and compares its resulting outputs against the desired outputs. Errors are then propagated back through the system, causing the system to adjust the weights which control the network. This process occurs over and over as the weights are continually tweaked. The set of data which enables the training is called the training set. During the training of a network the same set of data is processed many times as the connection weights are ever refined.
- **Unsupervised learning:** in unsupervised training, the network is provided with inputs but not with desired outputs. The system itself must then decide what features it will use to group the input data. This is often referred to as self-organization or adaption.

Although powerful, ANN are not without limitations. Taking into account the nature of this technique, the resulting applications experience a lack of ability to work with large numbers of irrelevant input variables. Furthermore, there can be some difficulty related to the interpretation and understanding of the extracted knowledge, since the internal functioning of the created models is considerably complex. Finally, in some cases, the computational time during training process can be very high due to a slow convergence of the learning procedure.

Support vector machines

The Study on Statistical Learning Theory started in the 1960s, having been under development by Vapnik (1998). Support vector machines (SVM) are a practical learning method based on Statistical Learning Theory. SVM have shown high learning capabilities even when working with complex data and can be used for either classification or regression analysis. For a given dataset, the SVM algorithm fits an unique and globally optimal solution. The underlying principle of SVM is to map the original data into a higher dimensional feature space and to optimally fit a linear function in this feature space.

SVM are a specific class of algorithms characterized by the use of kernels, absence of local minima during learning phase, sparseness of the solution and capacity control obtained by acting on the margin or number of support vectors. The nonlinear kernel functions are used to implicitly map inputs into high dimensional feature spaces (Figure 3.5).

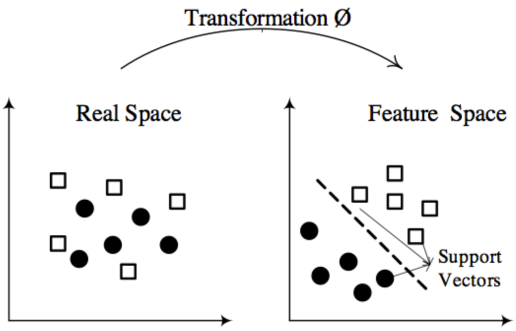


Figure 3.5 Example of SVM transformation (adapted from Cortez 2010b)

In this feature space, the SVM attempts to determine the best linear separating hyperplane. Some previously developed methods, such as the perceptron (Rosenblatt and Cornell Aeronautical Laboratory 1958), already used the concept of finding a separating hyperplane, but not the optimal one. SVM finds an optimal solution by maximizing the distance, or margin, ρ, between the hyperplane and the "difficult points" close to decision boundary (Figure 3.6). This is supported by the premise that if there are no points near the decision surface, then there are no uncertain classification decisions. It is noteworthy to add that the optimal dividing hyperplane is determined by a few parameters, namely by the support vectors, which correspond to the points closest to the hyperplane. Optimal separation of the support vectors is equivalent to optimal separation of the entire data.

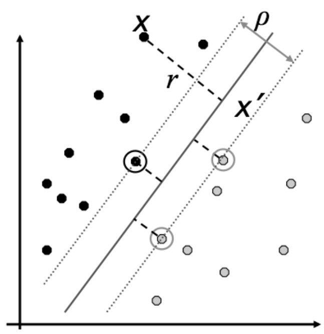


Figure 3.6 Example of hyperplane margin, ρ, and support vectors (circled points)

This method of representing decision functions is especially useful for a high dimensional input space: the number of free parameters in this representation is equal to the number of support vectors but does not depend on the dimensionality of the space (Vapnik et al 1997). Although SVMs are linear learning machines with respect to the feature space, they are in effect nonlinear in the original input space. This means that SVM can learn nonlinear behaviors without the drawbacks of nonlinear approaches, i.e., occurrence of local minima convergence problems. The popularity of SVMs is mainly due to their capacity to combine the advantages of linear and nonlinear models, as well as their predictive results that were achieved in several domains.

3.2.2. Model assessment

Metrics

After the training of a model using the training set of data, it is usually applied to a set of test data in order to validate and assess its learning capabilities. Depending on the type of problem being solved, classification or regression, different evaluation measures, or metrics, can be applied. In regression, metrics are based on the output error, defined as the difference between observed and predicted values.

The assessment of models resulting from the application of DM in this work was primarily based on the value of the error defining the degree of learning of a given model, as well as the correlation between the observed and the predicted values (Hastie et al 2009). The two mainly used metrics were the root mean squared error (RMSE) and the correlation coefficient (R^2) (Equation 3.3). Low values of RMSE, as well as R^2 values close to 1 should be interpreted as high model predictive capacity. The main difference between RMSE and mean absolute deviation (MAD) is that the former is more sensitive to extreme values since it uses the square of the distance between the real and predicted values. When compared with MAD, RMSE penalizes more heavily a model that in a few cases produces high errors. Thus, these were considered more appropriate for the applications of DM models discussed in Chapter 5.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y-\hat{y})^2}{N}} \; ; \qquad R^2 = \left(\frac{\sum_{i=1}^{N}(y-\bar{y})\times(\hat{y}-\bar{\hat{y}})}{\sqrt{\sum_{i=1}^{N}(y-\bar{y})^2\times\sum_{i=1}^{N}(\hat{y}-\bar{\hat{y}})^2}}\right)^2 \tag{3.3}$$

where

y - is the computed network output vector,

$\hat{y}$ - is the target output vector, and

N - is the number of samples in the database.

Generalization capacity

One of the main issues in model evaluation is to find the best way for the system to learn the concept represented by the training set, instead of learning the training set itself, that is, to achieve good generalization capacity. In other words, the generalization capacity of a DM model corresponds to how well it is able to accurately predict unseen values. The most common methods to infer on the generalization capacity of a predictive model are holdout, cross-validation and leave-one-out. In order to improve model reliability, each one of these approaches can be performed several times (executions, also known as runs).

The holdout method reserves a certain amount of data (from a dataset) for testing, while using the remaining data for training. Since the main purpose of the training set is to induce the model, 2/3 of the available data is often allocated to it, leaving the remaining 1/3 for model accuracy assessment. Although this method is advantageous in terms of simplicity and speed, the samples taken may not be very representative of the whole data, potentially resulting in different results for different random data splits.

Cross-validation, or k-fold validation, is an upgrade of the holdout approach, permitting the use all the available data for both training and testing (Figure 3.7). The first step in this method is the separation of data into k subsets of equal size. Secondly, each subset is used in turn for testing, while the remaining subsets are reserved for training. The error of the model is taken as an average from the errors of each iteration. Typical values for k are 5, 10 or 20, depending on the dimension of the dataset. Although being more robust than holdout, this method naturally demands a higher computational effort.

The leave-one-out approach (Hastie et al 2009) may be considered a special case of cross-validation, in which the number of folds corresponds to the number of training instances. In other words, in each interation only one example in the dataset is used for test, while the remaining data is used for training. This method is especially suited for small datasets (i.e., lower than 100 examples), since it can be computationally expensive. However, it involves no random subsampling, subsequently making the best possible use of the available data. The final generalization estimate is evaluated by computing evaluation metrics for all test samples.

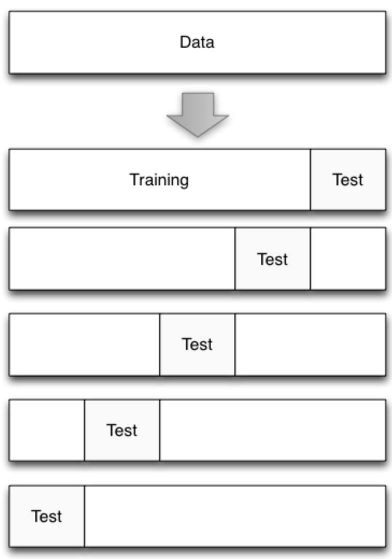


Figure 3.7 Cross-validation approach

3.3. METAHEURISTICS

The concept of artificial intelligence is far from limited to Machine Learning algorithms, since it includes several other applications, as is the case of metaheuristics, also included in the soft computing concept. In general, any task can be perceived as a problem for which there is a certain space of potential solutions. Considering that, in most cases, the main goal of solving a problem is the determination of the best solution, these problems can be seen as optimization problems. Even though, the classical methods of exhaustive search for solutions may be enough in a relatively small solution space, in many cases the solution space is considerably larger, demanding the need of different search methods based on modern heuristics (Michalewicz 1999). In the context of earthworks, most successful applications are based on genetic algorithms and swarm intelligence techniques, while a few have been developed exploiting fuzzy logic and Petri nets. Accordingly, the former techniques, representing the majority of successful earthwork optimization applications, will be given a more detailed description on account of their increased relevance.

3.3.1. Genetic algorithms

A genetic algorithm (GA) (Holland 1975) is a stochastic search technique used in computing to find true or approximate solutions to optimization and search problems. GA are categorized as global search heuristics. They consist of a particular class of evolutionary algorithms that use techniques inspired by evolutionary biology such as inheritance, mutation, selection, and crossover (also known as recombination).

GA are implemented as a computer simulation in which a population of abstract representations (called chromosomes or the genotype or the genome) of candidate solutions (called individuals, creatures, or phenotypes) to an optimization problem evolves toward better solutions. Traditionally, solutions are represented in binary as strings of 0s and 1s, but other encodings are also possible. The evolution usually starts from a population of randomly generated individuals and occurs over generations. In each generation, the fitness of every individual in the population is evaluated, multiple individuals are selected from the current population (based on their fitness), and modified (via crossover and possibly mutation) to form a new population. The new population is then used in the next iteration of the algorithm. Commonly, the algorithm terminates when either a maximum number of generations has been produced, or a satisfactory fitness level has been reached for the population. If the algorithm has terminated due to a maximum number of generations, a satisfactory solution may or may not have been reached.

Thus, the vocabulary used when dealing with GAs is similar to that used to refer phenomena of genetic evolution:

- **Individual** - Any possible solution;
- **Population** - Group of all *individuals*;
- **Search Space** - All possible solutions to the problem;
- **Chromosome** - Blueprint for an *individual*;

- **Trait** - Possible aspect (*features)* of an *individual*;
- **Allele** - Possible settings of trait (black, blond, etc.);
- **Locus** - The position of a *gene* on the *chromosome*;
- **Genome** - Collection of all *chromosomes* for an *individual*.

A typical genetic algorithm requires two things to be defined:

- a genetic representation of the solution domain; and
- a fitness function to evaluate the solution domain.

A standard representation of the solution is as an array of bits. Arrays of other types and structures can be used essentially in the same way. The main property that makes these genetic representations convenient is that their parts are easily aligned due to their fixed size, facilitating simple crossover operation. Variable length representations may also be used, but crossover implementation is more complex in this case. Examples of chromossomes are shown in Table 3.3.

Table 3.3 Chromossome types and examples

Type	***Example***
Bit strings	(0101 ... 1100)
Real numbers	(43.2 -33.1 ... 0.0 89.2)
Permutations of element	(E11 E3 E7 ... E1 E15)
Lists of rules	(R1 R2 R3 ... R22 R23)
Program elements	(genetic programming)
Any data structure	…

The fitness function is defined over the genetic representation and measures the *quality* of the represented solution. The fitness function is always problem dependent. For instance, considering the knapsack problem the aim is to maximize the total value of objects that can be put in a knapsack of some fixed capacity. A representation of a solution might be an array of bits, where each bit represents a different object, and the value of the bit (0 or 1) represents whether or not the object is in the knapsack. Not every such representation is valid, as the size of objects may exceed the capacity of the knapsack. The *fitness* of the solution corresponds to the sum of values of all objects contained in the knapsack if the representation is valid, or 0 otherwise. In some problems, it is hard or even impossible to define the fitness expression; in these cases, interactive genetic algorithms are used. The most common type of genetic algorithm follows the steps described in detail below (Figure 3.8):

- A population is created with a group of randomly generated individuals.

- The individuals in the population are then evaluated.
- The evaluation function is provided by the programmer and gives the individuals a score based on how well they perform at the given task.
- Two individuals are then selected based on their fitness, the higher the fitness, the higher the chance of being selected.
- These individuals then "reproduce" to create one or more offspring, after which the offspring are mutated randomly.
- This continues until a suitable solution has been found or a certain number of generations have passed, depending on the needs of the programmer.

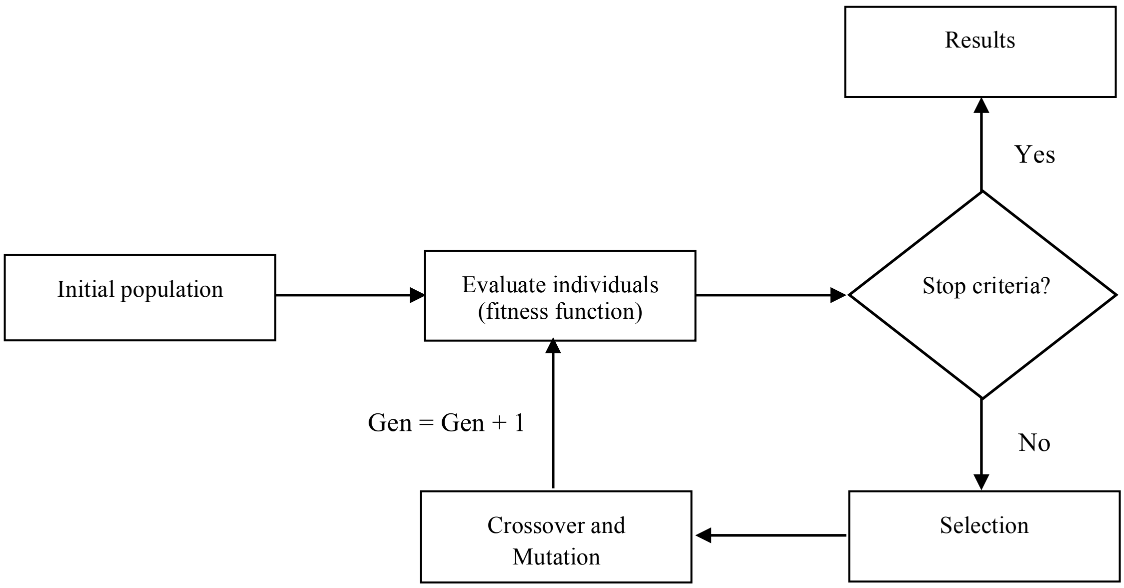


Figure 3.8 General algorithmic flow for GA

Initialization

Initially many individual solutions are randomly generated to form an initial population. The population size depends on the nature of the problem, but typically it contains several hundreds or thousands of possible solutions. The population is generated randomly, covering the entire range of possible solutions (search space). Occasionally, the solutions may be "seeded" in areas where optimal solutions are likely to be found.

Selection

During each successive generation, a proportion of the existing population is selected to breed a new generation. Individual solutions are selected through a fitness-based process, where fitter solutions (as measured by a fitness function) are typically more likely to be selected. Certain selection methods rate the fitness of each solution and preferentially select the best solutions. Other methods rate only a random sample of the population, as this process may be very time-consuming. Most functions are stochastic and designed so that a small proportion of less fit solutions are selected. This helps keep the diversity of the population large, preventing premature convergence on poor solutions. Popular and well-studied selection methods include roulette wheel selection and tournament selection. In roulette wheel selection, individuals are given a probability of being selected that is directly proportionate to their fitness. Two individuals are then chosen randomly based on these probabilities and produce offspring.

Reproduction

The next step is to generate a second generation population of solutions from those selected through genetic operators:

- **Crossover**: the most common type is single point crossover. In single point crossover, a locus is chosen at which the remaining alleles are exchanged from one parent to the other (Figure 3.9). The point at which the chromosome is broken depends on the randomly selected crossover point, considering a single point crossover, or points, in the case of multiple point crossover. However, crossover does not always occur. Depending on a set probability, crossover may be skipped, resulting in the parents being copied directly to the new population. The probability of crossover occurring is usually between 60% to 70%.

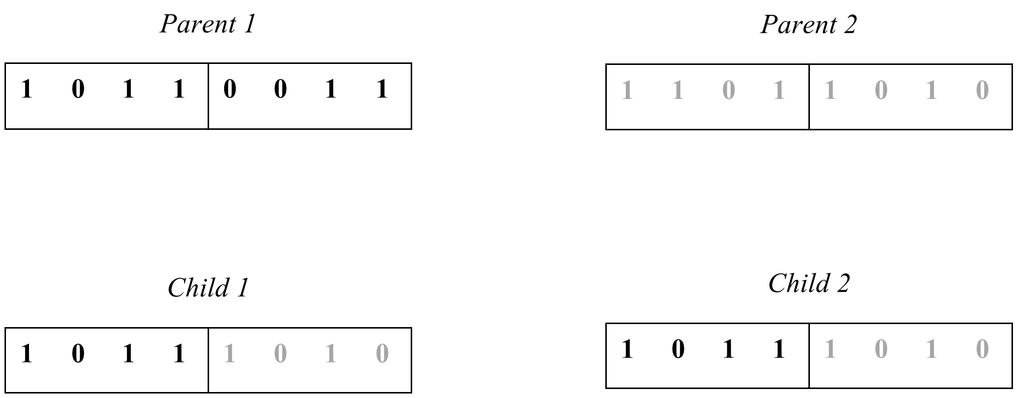


Figure 3.9 Example of single point crossover on binary chromossomes

- **Mutation**: After selection and crossover, a new population full of individuals is available. In order to ensure that the individuals are not all exactly the same, a small chance of mutation must be allowed. If that allele in an individual is selected for mutation, it can either change it by a small amount or replace it with a new value (Figure 3.10). The probability of mutation is usually between 1 and 2 tenths of a percent. Mutation is vital to ensuring genetic diversity within the population and allowing the appearance of really innovative features.

Before	**1**	**0**	**1**	**1**	**1**	**0**	**1**	**0**
After	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**

Figure 3.10 Example of mutation on binary chromossomes

For each new solution to be produced, a pair of "parent" solutions is selected for breeding from the pool selected previously. By producing a "child" solution using the above methods of crossover and mutation, a new solution is created, which typically shares many of the characteristics of its "parents". New parents are selected for each child, and the process continues until a new population of solutions of appropriate size is generated. These processes ultimately result in the next generation population of chromosomes that is different from the initial generation. Generally, the average population fitness will have increased by this procedure, since only the best organisms from the first generation are selected for breeding, along with a small proportion of less fit solutions, for reasons already mentioned above.

Termination

This generational process is repeated until a termination condition has been reached. Common terminating conditions are:

- A solution is found that satisfies minimum criteria;
- Fixed number of generations reached;
- Allocated budget (computation time/money) reached;
- The highest ranking solution's fitness is reaching or has reached a plateau such that successive iterations no longer produce better results;
- Manual inspection; and
- Any Combinations of the above.

3.3.2. Swarm intelligence

Swarm intelligence (SI) is the property of a system whereby the collective behaviors of (unsophisticated) agents interacting locally with their environment cause coherent functional global patterns to emerge. SI provides a basis with which it is possible to explore collective (or distributed) problem-solving without centralized control or the provision of a global model. A swarm resorts to the power of complex adaptive systems to solve difficult non-linear stochastic problems, displaying the following characteristics:

- Distributed, no central control or data source;
- Limited communication;
- No (explicit) model of the environment;

- Perception of environment (sensing); and
- Ability to react to environment changes.

Social interactions (locally shared knowledge) provides the basis for unguided problem-solving, where the efficiency of the effort is related to, but not dependent upon, the degree or connectedness of the network and the number of interacting agents. This type of problem-solving can be found in Nature in situations where survival in hostile environment, by means of adaptation and social interaction, is essential (e.g., bacteria, immune system, ants, birds and other social animmals). Although there are several models and algorithms that can be connected to the SI area, particle swarm optimization (PSO) and ant colony optimization (ACO) stand out as the main and most widely used methods.

Introduced by Kennedy & Eberhart (1995), the underlying concept of PSO takes its inspiration from the social behaviour of bird flocking or fish schooling. It is a population-based stochastic optimization technique for optimizing nonlinear functions using a particle swarm methodology. Swarms typically follow a specific behaviour:

- Steer toward the center;
- Match velocity of neighbors; and
- Avoid collisions.

For instance, consider a situation where a group of birds are randomly searching food in an area. There is only one piece of food in the area being searched. None of the birds know where the food is, but they know how far the food is in each iteration. In this situation, the most effective strategy to find food is to follow the bird which is nearest to it.

In PSO, each single solution is a "bird" in the search space, named particle. All particles have fitness values, which are evaluated by the fitness function to be optimized, as well as velocities, which direct the flying of the particles. During flight, each particle adjusts its position according to its own experience, and according to the experience of a neighboring particle, making use of the best position encountered by itself and its neighbor.

The algorithm initializes with randomly generated particles, going through several iterations in search for optima. Each particle has a velocity and position, which is updated every iteration by using two best values:

- Pbest: best solution (fitness) it has achieved so far; and
- Gbest: best global value obtained so far by any particle in the population.

In comparison with other algorithms, such as GA, one could say that both PSO and GA are population based stochastics optimization algorithms, starting with a random population which is successively evaluated by a fitness function. Moreover, like GA and other metaheuristcs, PSO makes few or no assumptions about the problem being optimized, being able to search very large spaces of potential solutions, but not guaranteeing the single best solution.

However, PSO does not have genetic operators, such as crossover and mutation, rather relying on updating the internal velocity of particles. The information sharing mechanism in PSO is significantly different, as it flows from best particles to the others, unlike GA, in which the population moves together. Finally, PSO is the only evolutionary algorithm that does not remove candidate population members, since all particles survive for the length of the run, disregarding selection steps.

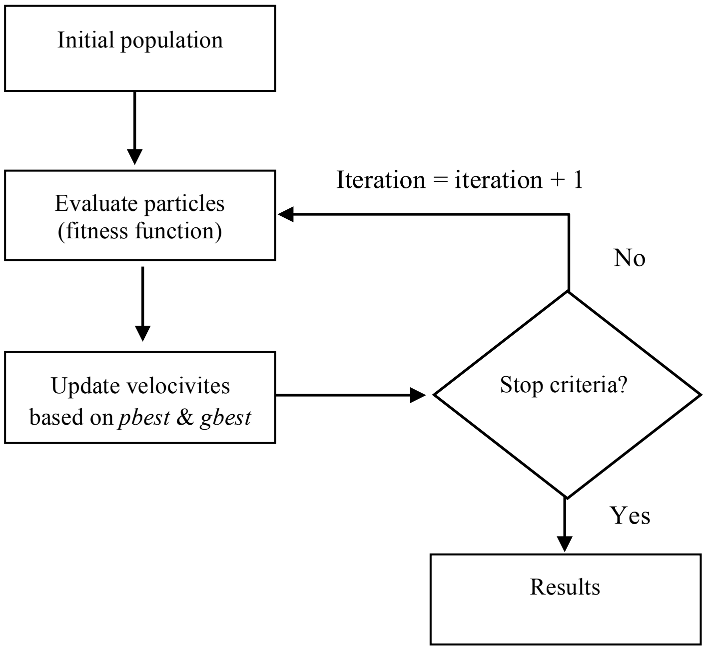


Figure 3.11 General algorithmic flow for PSO

Instead of relying on the swarm behaviour of bird flocks or fish schools, ACO are based on the behaviour of ant colonies and their natural adaptations for survival. Initially proposed by Dorigo and Gambardella (1997), these are agent-based systems which, simulating the natural behaviour of ants, develop mechanisms of learning and cooperation, being especially efficient for path finding (i.e., the traveling salesman problem) and combinatorial optimization problems. An ACO uses artificial "ants" to find paths in the search space. "Ants" leave trails of "pheromones" as they randomly go through different paths between their colony, or nest, and the food source. Those trails become stronger in shorter paths, where ants travel faster and more regularly, thus corresponding to better solutions. Paths with stronger pheromone trails have higher probability of being preferred by more ants. An evaporation rule is tied to the pheromones, which reduces the chance of poor quality solutions being chosen.

Figure 3.12 illustrates the process of foraging by ants, a clear example of how both nature and ACO carry out optimization in terms of finding the best path. When ants leave their colony, or nest, to search for a food source, they randomly rotate around any obstacle in their way, as exemplified in Figure 3.12a. Initially, the pheromone deposits will be the same for both illustrated paths (Figure 3.12b). At the time when the ants find a food source, they carry the food all the way back to their nest, following their own pheromone trails and still depositing more pheromone on their corresponding trail (Figure 3.12c). An ant that is looking for food is more likely to choose the shortest path when leaving the nest, knowing that this will be the path with the most deposited pheromone, since the ants there travel quicker and less evaporation of pheromone is verified (Figure 3.12d). For this reason, new ants that later start out from the nest to find food will also choose the shortest path, increasing the pheromone level on that path even further (Figure 3.12e). Over time, this positive feedback process prompts all ants to choose the shortest known path, as depicted in Figure 3.12f.

In the previous example, only one node was included in the path finding problem. However, the algorithm can deal with path finding under situations with several nodes. Upon reaching a node for the first time, ants choose a path randomly. However, as time progresses and pheromones are deposited, as an ant reaches a node, it will have a higher probability of following through a path with more pheromones. This process is repeated for as many cycles as necessary until most ants select the same tour on every cycle (convergence criteria).

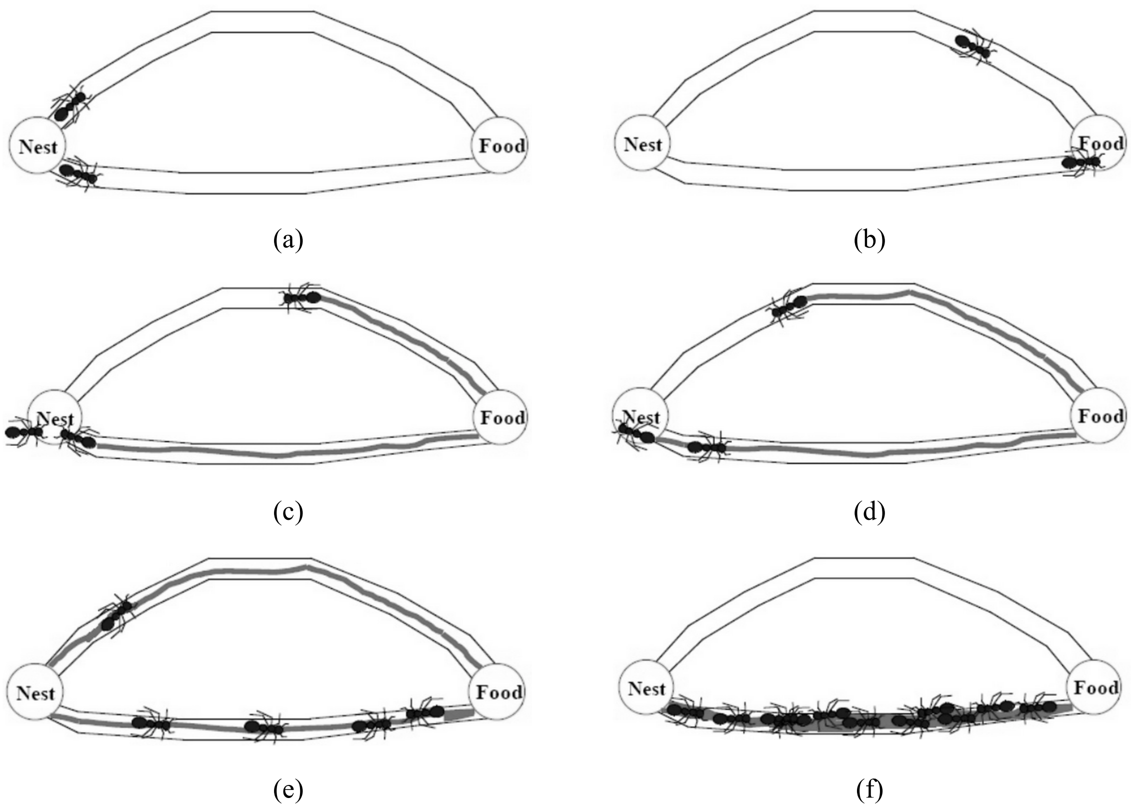


Figure 3.12 Ant foraging – co-operative search by pheromone trails

The pseudocode for an ACO algorithm can be defined as follows:

- Initialize Trail
- **Do While** (Stopping Criteria Not Satisfied) – Cycle Loop
    - **Do Until** (Each Ant Completes a Tour) – Tour Loop

  - Local Trail Update
  - **End Do**
  - Analyze Tours
  - Global Trail Update
- **End Do**

As with other evolutionary algorithms such as GA, ACO is a good choice for constrained discrete problems, as it does not base its search on gradient search methods. Its adaptability to changes (e.g., new distances, restrictions or obstacles) makes it ideal to be used in dynamic applications. The fact that ACO algorithms retain memory of the entire colony (rather than of the previous generation only), and that it is less affected by poor initial solutions (due to combination of random path selection and colony memory) gives it an edge in some problems when compared to evolutionary algorithms.

Yet, ACO algorithms also feature a few disavantages in comparison with other metaheuristics, mainly in terms of convergence and interpretation. Firstly, due to its nature, while convergence in ACO algorithms is guaranteed, the time to convergence is always uncertain due to the randomness implied in the algorithm. Furthermore, a large number of different ACO algorithms are required to exploit different problem characteristics. This is worsened by the fact that the coding of ACO is somewhat complex, due to aspects like pheoromone evaporation and global and local updates. Finally, the sequences of random decisions taken by ants, as well as the changes in probability distribution in every iteration/cycle, can also hinder the interpretation and theoretical analysis of results.

3.3.3. Fuzzy logic and Petri nets

Fuzzy logic algorithms (FLA) and Petri nets (P/T nets) have also been successful in terms of known earthwork optimization applications. On the one hand, fuzzy logic was conceived by Zadeh (1965) and consists of a mathematical technique for dealing with imprecise and vague data/knowledge, as well as problems that have many solutions rather than one. Inspired by human decision-making, it represents a type of logic that recognizes more than simple true and false values, since propositions can be represented with degrees of truthfulness and falsehood. For example, the statement "today is sunny" might be 100% true if there are no clouds, 80% true if there are a few clouds, 50% true if it is hazy and 0% true if it rains all day. FLA are used for solving problems with expert systems and other intelligent systems that must react to an imperfect environment of highly variable, volatile or unpredictable conditions. Its advantage is its ability to deal with vague systems and its use of linguistic variables.

On the other hand, P/T nets (Petri 1966) are mathematical models based on bipartite graphs (i.e., graphs whose vertices can be divided into two disjoint sets, such that every edge connects a vertex in the first set to one in the second set), which constist of transitions (representing events), places (corresponding to conditions) and directed arcs (which signify the dependencies and relationships between places and transitions). On account of their nature, P/T nets are especially well suited for modelling the concurrent behaviour of dynamic systems. Conditions and relationships may be associated with operations like sequential execution, synchronization, merging, concurrency and conflict. Although P/T nets are not always associated with SC techniques, they have been known to be coupled with the latter for modelling or optimization purposes, thus justifying their mention in this work.

3.4. GEOGRAPHIC INFORMATION SYSTEMS

According to Denègre and Salgé (1996) and Gutiérrez Puebla and Gould (1994), a geographic information system (GIS) can be considered a specific type of information system. Indeed, an information system can be defined as a communication system, based on programmed routines or integrated subroutines, that is capable of manipulating information, such as large databases including extensive real-world data (Martin 1991). Under this concept, GIS focus on communicating and manipulating geographic information, dealing with real-world data. Its final goal is decision support, using tools for handling, representing and communicating geographic knowledge.

Geographic information is regarded as the representation of objects of real phenomena in a specific location at a given time. Overall, the concept of geographic information encompasses two main points:

- Information regarding one object of phenomenon, described in terms of nature, shape and attributes (e.g., a road can be described by its type, pavement, number of lanes, etc.). This information may include the description of relations with other objects or phenomena (e.g., a building belonging to a proprietor).
- Information regarding the location of the object, relatively to a specific reference system (e.g., coordinate system or postal address).

The origins of GIS can be traced to the computerization of cartographic information (i.e., “computer mapping”) recorded in the 1960s. This data was progressively compiled into several databases, within the concept of an information system, aimed at managing, archiving and updating this data. Through the 1980s the concept was already established and several commercial vendors of GIS software emerged, including capabilities such as:

- Managing and manipulation spatial relations between objects, using spatial analysis functions for decision support;
- Representing space and its components as a layout or map, implying the existence of cartographic construction functions.

In order to achieve this, a GIS should include at least the following data modules (Peuquet and Marble 1990):

- A subsystem for data entry, capable of data acquisition from existent maps, GPS emitters, etc.;
- A subsystem for data storage and retrieval, which organizes the spatial data so that it can be easily accessed by the user for subsequent analysis, as well as updated and altered;
- A subsystem for data manipulation and analysis, responsible for carrying out different tasks, namely aggregating the data according to user input, or estimate parameters from spatial simulation or optimization models;
- A subsystem for outputting and presenting information, for displaying the database of the processed data, as well as the results from spatial analysis.

Within this framework, a GIS must be able to provide the basic information pertaining **what** an object is, **where** it is located and **when** and **how** it relates with other objects. Furthermore, since it allows spatial analysis and manipulation, different **what if** scenarios and queries must also be provided by the system.

3.4.1. GIS components

According to Zeiler (1999), a GIS is the combination of experts, spatial and descriptive data, analytic methods, software and hardware, aimed at managing and displaying geographic information. This description encompasses five essential components that comprise a GIS, namely people (liveware), hardware, software, data and methods (Figure 3.13).

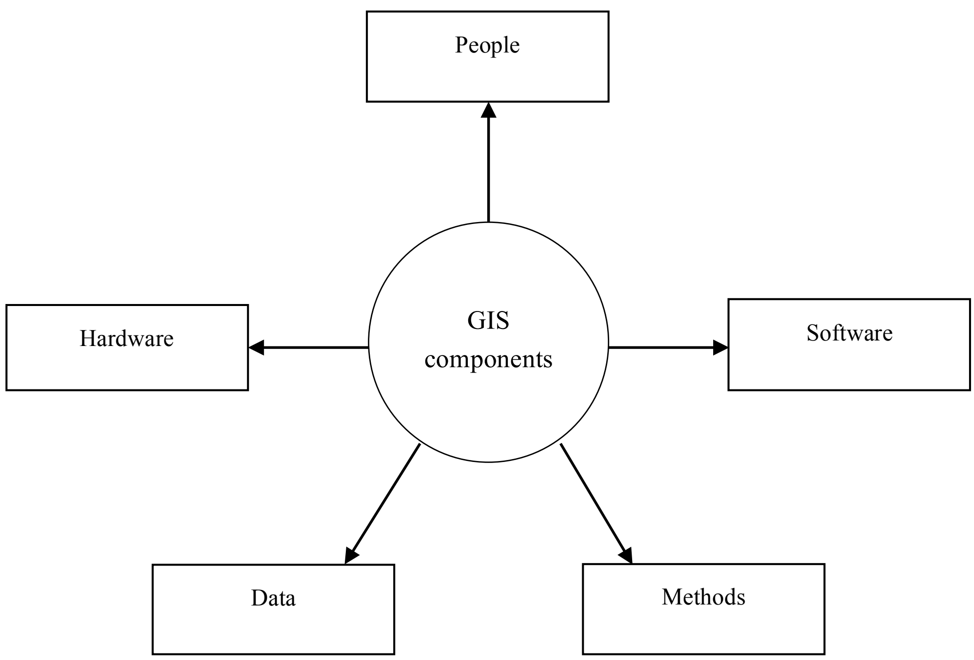


Figure 3.13 Components of a GIS (adapted from Zeiler 1999)

The hardware is the physical component of a GIS, acting as the interface between the human (people) and the software components. On the one hand, it is a platform (e.g., from a personal computer to powerful working stations) for the GIS software installation, data storage and communication (e.g., network or internet access). On the other hand, it represents a means for acquiring data (e.g., by supporting the use of scanners and similar equipment), as well as displaying results (e.g., by adding printers, plotters, etc.).

Being dependent on proper hardware to support its installation, the software is the component that allows for data manipulation and analysis. With the development of GIS, a large variety of commercial software has emerged, though with similar objectives and functions. The main differences among them are often related to how the data is stored, which type of uncommon operations they can perform, their implementation methodologies, user-friendliness and user support.

Data are an abstract representation and, in general, simplification of the real world. They are often the most critical part of the GIS, since operations and analysis methods are carried out on them and the usefulness, as well as reliability, of their output is directly related to the quality of the data. According to Rowley and Gilbert (1989), a set of quality data can amount up to 70% of the total cost of a project, which emphasizes the importance of this component.

The analysis methods and objectives performed on the data depend on the expertise and intentions of the user. Meteorology, hydrology, epidemiology are a few examples of scientific areas which make extensive use of GIS software. Functionalities can range from procedures that guarantee the quality of the data to employing algorithms, to solving spatial problems in linear or polygonal networks, or to carrying out tridimensional visualisations of areas using cartographic data.

Finally, the people, or experts, represent the human component and are the ones who, using the hardware as a support for software, perform spatial analysis and manipulation on the available data. Whatever the purpose of the analysis (e.g., urban planning, infrastructure development), the user will be the one making the decisions according to the output of the system.

3.4.2. Data entry and acquisition

Considering data is one of the critical components of a GIS, the available methods to add or acquire data into are paramount. The usual methods to achieve this are:

- Import digital information available in compatible formats;
- Use of global positioning system (GPS) apparatus; and
- Manual digitizing of analog data.

On the one hand, the development of GIS has been accompanied by an emergence of digital information and databases, to whose access is becoming easier and easier as communication technologies develop, especially the Internet. On the other hand, compatibility between software is increasingly common, which makes it possible to convert data originated in one software to be used in another. Formats like CAD (e.g., dwg, dxf), vectorial and raster data from some broadly used commercial GIS (ARC/Info, ARC/View, Intergraphe MGE, etc.), and general image data (e.g., tiff, bmp, etc.) are some examples of data than can be added directly into most GIS software.

Another widely known method for data acquisition using GIS is the use of global positioning system (GPS). By emitting a signal to several satellites and using triangulation to determine its position and altitude with low error margins (e.g, below one meter), GPS can be connected to a GIS for various purposes, such as mapping roads or knowing the location of trains, buses or construction equipment at any given time.

In cases where data is not available in any other format but analog, it is necessary to manually digitize it. This can be done by one of two ways: automatically, by means of a scanner, or manually, using a digitizing table. In the first case, one obtains a raster file that can be converted into vectorial data if required. In the second case, the obtained data is in vectorial format ready to be manipulated or analysed. After digitizing, it is still necessary to build the topology of the data, in which the system automatically adds extra information to the database, such as polygonal areas or arc lengths.

In cases where data entry is carried out by using CAD, image files (e.g., tiff, bmp), or manual digitizing, it is also required to add the thematic attributes, which allow for the determination of other variables that can then be used in spatial analysis. This ranges from georeferencing an image file, for instance by identifying the coordinates of known points in the image, allowing the system to infer the coordinates and distances of the remaining area (the more points are added, the more exact the inference process is), to identifying a potential point or vector of interest, in order to use it for subsequent analysis. This step must be carried out manually or by importing the information from other sources.

3.4.3. Network analysis

Among the various capabilities of GIS, network analysis is of special interest to this work and following Chapters. This type of analysis is limited to vectorial data, since its use implies the availability of a network, defined as a series of interconnected linear features, which represent the potential routes for circulation of goods, people, resources, services or information, depending on established constraints (Heywood et al 2011). One of the key features of any network is connectivity. In this point of view, correct geographical representation in a network may not be as important as the correct representation of connectivity and its associated characteristics. Accordingly, a network is usually comprised of three main elements (Figure 3.14):

- Links are the linear connection between two nodes, usually used to represent the routes;
- Nodes correspond to interceptions, as well as origin and destination points, depending on their associated type, which are depicted in Figure 3.14:
    - Vertices;
    - Junctions;

- End nodes;

- Turns represent the transition from one link to another.

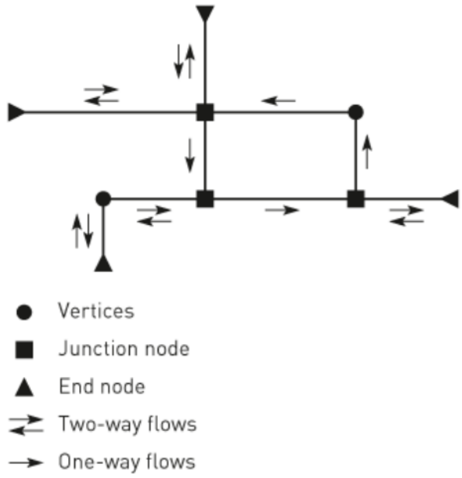


Figure 3.14 Network elements (Heywood et al 2011)

Transportation, electricity, water supply and sewer lines are some examples of common application of networks in GIS (Figure 3.15). The type of analysis can be widely different, placing emphasis on the routing problems, where the shortest route for transportation equipment, or other vehicles, can be determined in function of variables such as possible routes, maximum speed in each route or one-way restrictions. The routing problem will be further discussed in this Section, since it is relevant for the following Chapters.

Figure 3.15 Network examples

In the context of earthwork optimization, one of the most useful features provided by GIS software is the determination of minimum routes between two points in a network. In general, the shortest route between two nodes in a network corresponds to the minimum distance between nodes. However, in most cases, the factor that influences the determination of the shortest route is impedance rather than distance. Impedance is a measure of the amount of resistance, or cost, required to traverse a path in a network, or to move from one node to another. Impedance may be measured in terms of factors like time, distance and fuel usage, among others. Furthermore, it may be a function of aspects such as topography, traffic volume or weather. To sum up, impedance in a network attempts to reflect the conditions in which circulation is accomplished (Gutiérrez Puebla and Gould 1994) and can be assigned to both routes (i.e., links) and interceptions (i.e., nodes):

- Link impedance: represents the resistance to movement between the extremities of a link. The simplest measure of impedance for links is their length, since the longer the link, the higher the effort necessary to go through it. Other options include surface grade, type of surface/pavement, presence of obstacles, etc. Time is the most used measure for vehicle routing problems, since it is directly related to speed and distance, providing the flexibility for different types of analysis. It is noteworthy to add that impedance can also include a direction feature, allowing for analysis such as one-way routes or different speeds in each direction for a single route. Whichever the case, deciding on the type of impedance for each case is a prerequisite for spatial analysis.
- Node impedance: generally attempts to simulate the circumstances associated with changing direction at interceptions. For vehicle routing problems, they can be used, for instance, to simulate the difference in required time to turning left or right at interceptions, as a result of traffic. In a railway model, node impedance may represent the time that a train stops at each station. As such, the total impedance between two nodes in a network is equal to the sum of impedance of links and nodes for the shortest path between the nodes.

It is clear how the shortest route between two points in one city can be different, depending on the considered impedance: the shortest path (distance) may go through the city center, while the fastest path (time) will probably imply taking a highway in order to avoid the traffic in the city center. This example illustrates the increasingly high importance of this type of analysis in a wide range of activities, such as emergency vehicle routing, urban public transportation or commodity distribution. In the earthworks point of view, impedance can be used to represent a maximum velocity for each route and interception (i.e., links and nodes) in a haul road network, which, in turn, can be a function of factors such as road conditions, grade or rolling resistance. This would result in different haul cycles durations for different potential routes, one of which would correspond to the shortest path (i.e., lowest haul cycle duration).

In summary, in order to determine the minimum route between two points in a network, one must supply the location of points, the variable to consider as link impedance and, if applicable, the node impedance. Recalling this information, the GIS can output the minimum route through the network, identifying each the links, nodes and directions to take, as well as provide the value of the total impedance for the chosen route (Figure 3.16). Finally, it is relevant to refer that it is possible to assign more than one destiny, requesting the GIS to determine the shortest route that passes through all intermediate points. In this case, it is also possible to include information regarding supply and demand of cargo (goods or people) between the origin and destination points, making the GIS output the type and quantity of cargo to transport in each visit in order to maximize profit, as well as the order of intermediate destinations that minimizes impedance.

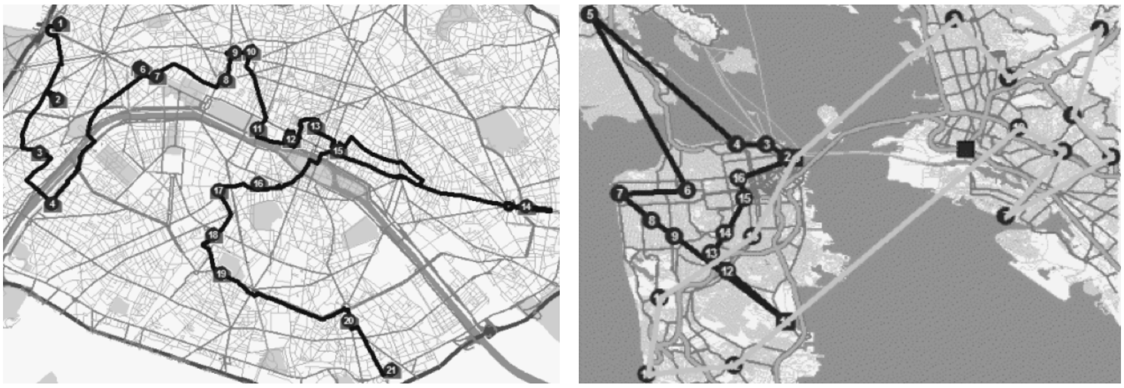


Figure 3.16 Routing examples

3.5. SOFTWARE AND TOOLS

3.5.1. R Tool

Achieving full integration of the mentioned technologies is to be underlined, considering the level of complexity to be dealt with by a system, which would aim at accomplishing goals as set in this work. Being able to handle the large amount of variables and situations that are verified in an earthwork construction by using a single system is not only practical in the point of view of the user, but also essential from a technical perspective. Indeed, the high sequentiality of earthwork tasks, allied with a wide range of relationships and constraints between available resources in each situation, calls for a high level of communication and interaction between the different technologies, justifying their integration in modules of a single system.

Nowadays, there are several data analysts that can be used to fulfil the data mining role, such as *RapidMiner*, *R*, *MS Excel*, *Weka*, *SAS* and *Matlab*, among others. From these, the commercial software *Matlab* and the open-source *R* (R Development Core Team 2011) feature the option to easily extends their capabilities by installing additional packages for different purposes, promoting the ideal environment for integration between data mining and optimization technologies, including metaheuristics. Furthermore, the software in question also includes the possibility of being integrated with geographic information systems, as discussed in the next section (Section 3.6.2).

In the specific context of this project, *R* shows the most potential for achieving the proposed objectives. On the one hand, previous development and experience with this software on the geotechnics area (Tinoco et al 2011; Gomes Correia et al 2012), and in particular in earthworks (Marques et al 2008), proves the potential and reliability of this programming language in the context of this project. On the other hand, the *R* environment is an open source, multiple platform (e.g. *Windows, Linux, Mac OS*) and high-level matrix programming language for statistical and data analysis (R Development Core Team 2011), whose features are consistent with the requirements of the project. In fact, *R* includes a variety of statistical (linear and nonlinear modelling, classical statistical tests, classification, clustering, etc.) and graphical tools, besides having a very flexible object-oriented design, becoming easily extensible by the installation of additional packages. The R community is highly active and new packages are continuously being developed, hence, in this perspective, the R can be seen as an open tool for worldwide sharing of algorithms. Furthermore, an extensive help system is included and available from the prompt, complemented by a large documentation freely available on the *R website* (http://www.r-project.org/), as well as on books (Muenchen and Hilbe 2010). As a drawback, the *R* tool has a relatively hard learning curve for non-expert users, due to the lack of a friendly graphical user interface (GUI), and the absence of technical support. Yet, after some experience and training, the user achieves a better control and understanding of what is being executed (in contrast with several "black-box" DM GUI products).

Regarding its DM potential, two of the most remarkable interfaces for *R* are the *Rattle* (Williams 2011) and *rminer* (Cortez 2010a) packages. While *Rattle* has the advantage of featuring a GUI, *rminer* is lighter and easier to install. Moreover, the latter presents more ANN and SVM capabilities and has been proven to perform better in many cases than other well-known DM tools (Cortez 2010b). Consequently, the *rminer* package is adopted in this work. This package is an integrated framework that facilitates the use of DM algorithms in classification and regression tasks, being particularly well suited for ANN and SVM algorithms, featuring a set of coherent and robust functions (Cortez 2010b; Cortez 2014):

- `fit` – create and adjust a given DM model using a dataset;
- `predict` – returns the predictions for new data;
- `mining` – a powerful function that trains and tests a particular model under several runs and a given validation method; and
- `mgraph`, `metric` and `mmetric` – return several mining graphs or classification/regression metrics.

Besides DM, the *R* environment also allows for the usage of optimization algorithms, including packages based on metaheuristics such as GA (Mersmann et al 2014; Willighagen 2015) and SI (Naval 2013). As discussed in more detail further ahead (Chapter 5), multi-criteria optimization is indicated for earthworks optimization. In point of fact, besides giving less priority to knowledge of optimization parameters, a Pareto approach outputs a set of solutions as a set of trade-offs between optimization objectives. This gifts the optimization process and its outcome with the versatility to handle the dynamic environment inherent to earthwork construction. As such, given the high success of genetic algorithms (as discussed in Chapter 4) and the versatility of Pareto approaches (further discussed in Chapter 5) for multi-criteria optimization, the package *mco* (Mersmann et al 2014) was adopted for the optimization process in this work. The package implements a non-dominated sorting genetic algorithm-II (NSGA-II) (Deb et al 2002), an evolutionary algorithm variant specifically designed for multi-objective optimization, founded upon three concepts that set it apart: Pareto optimization, elitism and sparsity. The function `nsga2` corresponds to the NSGA-II implementation in the *mco* package, featuring the following main parameters (Cortez 2014; Mersmann et al 2014):

- `fn` – function to be minimized (should return a vector with the several objective values);
- `idim` – input dimension;
- `odim` – output dimension (number of objective functions);
- … – extra arguments to be passed fn;
- `lower.bounds`, `upper.bounds` – lower and upper bounds;
- `popsize` – population size (default is 100);
- `generations` – number of generations (default is 100) or a vector;
- `cprob` – crossover probability (default is 0.7); and
- `mprob` – mutation probability (default is 0.2)

This function returns a list with the final population (assuming `generations` is a number, otherwise a vector list is returned where the i-th element contains the population after `generations[i]` iterations) with the components:

- `$par` – population values;
- `$value` – matrix with the best objective values for the last population individuals; and
- `$pareto.optimal` – boolean vector that indicates which individuals from the last generation belog to the Pareto front.

3.5.2. Geographic Information Systems

ArcGIS (ESRI 2011) is a set of commercial software products, produced by ESRI, embodying a geographic information system (GIS). Alternatively, *Quantum GIS* (QGIS) (Quantum GIS Development Team 2015) is a broadly used open-source geographic information system and a reference in the free and open-source (FOSS) community for users, developers and supporters. Both software feature the ability to work with cartographic and geographic information, being mainly used for the creation and use of maps, geographic data compilation, analysis of mapped information, as well as sharing and discovering spatial knowledge using the cartographic and geographic information. The systems are built around geodatabases, which implies an object-oriented approach to storing and manipulating spatial data. Thus, a geodatabase is a vessel for concretizing data sets by joining the spatial characteristics of objects (e.g., mechanical resources) with their associated attributes. In addition, they may also contain topology information, being able to model the behavior of the spatial characteristics of objects, integrating rules on how they relate to each other.

These tools include several toolboxes with particular importance to this work, namely the *ArcGIS Network Analyst* and the *QGIS Road Graph plugin*. The toolboxes include the capacity to build, analyse and manipulate network-based spatial data, allowing for a realistic modeling of dynamic conditions in a network, including one-way streets, turn and height restrictions, speed limits, and speed variation based on traffic. Among the various capabilities of the *Network Analyst*, those with the greatest potential for achieving the objectives proposed in this paper are emphasized:

- Generate a network from existing spatial data;
- Determine the most efficient path for a vehicle or a fleet of vehicles that must pass through multiple locations (i.e., routing and traveling salesman problems);
- Set time windows to limit when vehicles can reach specific locations;
- Determine ideal locations for facilities, performing location-allocation analysis;
- Generate origin-destination (OD) cost matrixes, from each starting point (e.g., excavation front) in the network to all associated destinations (e.g., embankment fronts).

Although featuring most of the same capabilities of *ArcGIS Network Analyst*, the *QGIS Road Graph plug-in* does not have a direct interface function for the generation of origin-destination cost matrixes. However, this limitation can be surpassed by programming a series of loops which, using the feature to solve simple routing problems from one point to another, carry out the necessary runs to determine the cost (e.g., in terms of distance or travel duration) associated with travelling from each source to each destination. By saving the results in the form of a matrix, this process results in an OD cost matrix. This is increasingly important considering that QGIS developers claim the possibility of integration with the R environment, meaning that it is possible to call commands from the R console directly to QGIS, as well as import the results of spatial analysis automatically for further usage in R. However, experimentation with this possibility revealed that integration is hindered by incompatibilities between the versions of QGIS that can be integrated with versions of R, as well as between the versions of R that allow integration and the data mining and optimization packages (i.e., rminer, mco) used to develop the algorithms. Nonetheless, there are some indications that the developers of both software are working to facilitate this integration in future versions. An alternative to this setup could be explored by integrating *ArcGIS* with the commercial object-oriented programming software, *Matlab*, with the obvious advantage that, a higher level of customer support could be expected, since both correspond to commercial software. Yet, bearing in mind that the R environment was the chosen foundation for the development of the optimization system for the reasons mentioned in the previous section, this work was initially directed towards using *QGIS* and R, despite the risk that limitations concerning the integration of both software could be extended indefinitely.

Google Earth (Google Inc. 2009) is a free software tool, developed and distributed by the company Google, whose function is to present a three-dimensional model of the globe, constructed from satellite pictures/aerial images (photographed from aircraft) obtained from various sources, as well as 3D GIS data. Thus, the program can be used simply to generate two-dimensional maps and aerial/satellite images or as a simulator of the wide range of landscapes existent on Earth. This makes it possible to identify areas, buildings, cities and landscapes, among other elements. Formerly known as Earth Viewer, *Google Earth* was developed by Keyhole, Inc, a company that Google acquired in 2004. The product name was changed in 2005 and is currently available for use on personal computers in various operating systems. Google made customer improvements to the Keyhole client and added a satellite database of images for its mapping software based on the World Wide Web. Most large cities on the planet are now available for images with enough resolution to view buildings, houses or even closer details as cars. The entire globe is already covered with an accuracy of at least 15 kilometers. In the earthwork construction context, *Google Earth* can be used as a source for aerial images of an area corresponding to a present or future site construction, as well as the corresponding coordinates for proper georeferencing. This information can be imported into any of the previously mentioned GIS software, allowing for different types of spatial analysis, such as determining optimal routes for transportation equipment throughout the construction site (e.g., via generation of OD cost matrixes).

3.6. FINAL REMARKS

In order to achieve the goals proposed in Chapters 1 and 2, several technologies are required to deal with the complex and dynamic environment that characterizes earthwork constructions. Given the sequential nature and high amount of relationships between variables of the earthowks optimization problem, the integration of these technologies into a single system has numerous advantages. In this context, the open source, high level programming langague *R* and associated software shows the most potential for the implementation of such system.

The resulting integrated system is presented in Chapter 5, following the discussion and analysis of the literature research on similar systems, with focus on the technologies they are based on, as well as their strengrths and limitations. This latter aspect is of critical importance in the study, as the key to innovation, lies in surpassing its predecessors' weaknesses, while encompassing their strong points.

Chapter 4

EARTHWORK OPTIMIZATION APPLICATIONS

4.1. INTRODUCTION

Decision-making in complex and dynamic environments such as Civil Engineering, often strains one's cognitive capabilities, especially when potentially large impacts are associated with the quality of these decisions. Aiming to support human judgment and decision-making, disciplines such as statistic, economics and production have researched various choice-supporting methodologies. When combined with techniques derived from information systems and artificial intelligence (AI), these methodologies can be integrated into computer programs or routines, thus originating what is commonly referred to as Decision Support Systems (DSSs) (Drudzel and Flynn 2002). The term Intelligent Decision Support System (IDSS), introduced by Holsapple (1977), is associated with DSSs that specifically make use of AI techniques. Conceptually, IDSSs are meant to be used as though they represent a human consultant, since the AI techniques, which are part of their core, intend to emulate human capabilities as closely as possible. These support the decision makers by gathering and analysing data and identifying problems, while proposing and evaluating potential solutions for the possible scenarios. Many IDSS implementation are based on Expert Systems (Sol 1985; Farinha et al 1995; Portela and Bento 2001), which encode the cognitive behaviour of human experts using logic rules, and have been known to be successfully applied to several different areas, including engineering (Marques et al 2008). The relation between DSSs and the concept of Business Intelligence (BI) is also worth highlighting, as some authors state it as the successor of DSSs (Power 2002; Turban et al 2007). Their purpose is to aid decision-making by extracting data (i.e., by means of data mining techniques) from different sources in order to support the potential choices. By presenting the ability to retrieve knowledge from large organized databases, symbolically representing heuristic knowledge and manipulating these in order to emulate human reasoning, AI approaches come forward with new and innovative problem-solving potential. Yet, the most obvious drawback in earthwork optimization context is the need for relatively large training dataset of past and relevant data or examples, so as to allow the models to extract knowledge and go through a "learning" process on the subject. Only then can this type of systems achieve the ability to manipulate this knowledge in the direction of decision support, ultimately finding a solution, or set of solutions, for a problem.

Unlike these, typical optimization models do not depend on learning from past data, but on logically built models that aim to find a solution for a problem with regard to a specific objective. In general, their structure consists of the definition of an objective function, for which a set of inter-related decision variables are manipulated in function of a set of constraints. Solutions can be found by means of different searching methods in the feasible space imposed by the set of constraints. Such methods often attempt to take advantage of the shape of the search space under the assumption that, when optimizing a problem, the optimal solution generally coincides with a maximum or a minimum point in the space created by those constraints. In the context of engineering applications, the advantages of optimization models include their efficiency, a simple understanding of the functioning of the solution searching methods and the easiness with which it is possible to evaluate the obtained outcome. However, pure optimization models are not without their downsides. Often they imply some level of unrealistic modelling assumptions and are only applicable to quantifiable problems, requiring specialized skills for both model formulation or alterations and solution interpretation.

Moreover, the choice of solution methods has to take into consideration not only the problem type, but also the risks of the solution getting stuck at local optima, which may represent limitations to model development (Dutta 1996).

Considering both AI and optimization technologies, one can easily infer that most limitations of one are balanced by the strengths of the other in the context of decision support and problem-solving. Optimization models are indicated for dealing with the structured and quantifiable aspects of the decision-making process, while AI techniques are directed for the more qualitative features of the process, which are usually performed by human experts. Thus, integrating the complementary strengths of AI and optimization seems to be a valid way to increase the decision support capabilities of systems, as discussed further on.

4.2. RESEARCH METHODOLOGY

In order to grasp the current state of development of intelligent earthwork optimization systems, one must examine, in detail, the different methodologies and system architectures that have been used to address this issue. Traditionally, the planning, construction and management of earthworks tasks are mostly based on the accumulated experience of experts or otherwise partially compiled in country-specific guides, as is the case of the *Guide des Terrassements Routiers* (GTR) compaction guide (SETRA and LCPC 2000). This is carried out under the premise that earthworks can be perceived as a dynamic production line based on heavy machinery. In this context, the work rate of the last process (compaction) determines the development rate of the whole construction, while simultaneously being dependant of the work rate of every preceding task. Hence, tools capable of guaranteeing a near optimum level of global productivity, while at the same time compensating for any lack of experience from the experts are needed. Yet, most research involving this subject focuses on supporting specific tasks or parts of the earthwork process, and therefore do not fully account for the interdependencies between processes in the context of a production line.

This section outlines the research methodology of this study, as well as the results from the literature search. The objectives were to compile information on the existent technologies and systems that, incorporating the technologies discussed in Chapter 3, have the potential to be part of earthwork optimization systems, analysing their strengths and limitations within this framework. Note that while this Section 4.3 fundamentally enumerates the relevant developments regarding the application of soft computing (SC) techniques in the earthwork optimization context, Section 4.4 contains the analysis of these systems, including their architecture and applications. In order to convey a deeper understanding of the functioning of the enumerated systems in Section 4.3, two systems (of different types) were selected for a more detailed description. The selection criterion was based on their relevance pertaining the remaining work. The resulting knowledge stemming from this Chapter was then put to use in developing the architecture of an intelligent earthwork optimization system, presented in Chapter 5, which can overcome the limitations identified in the analysed systems.

The research was conducted in three phases. Initially, the research goal and scope were defined. The former consists of the compilation the existent information regarding intelligent earthwork optimization systems, or in other words earthwork optimization systems that make use of SC techniques, while the latter consists of academic research from 1999 to 2013. As this topic is relatively recent, this 15-year period is deemed to be representative of the development and application of soft computing in earthwork optimization systems. The second phase consisted on the definition of the search criteria and the compilation of information gathered in books, journal and magazine articles, proceedings and technical reports. The search was based on the descriptors “earthworks” and “optimization”, resulting in an initial database of around 400 articles. Topic filtering reduced this number to 95 articles related to descriptors such as “soft computing” and “artificial intelligence”. Finally, a careful analysis of the remaining articles settled the search on 28 relevant systems and applications in this area. Master dissertations, doctoral thesis and unpublished working papers were excluded mainly on the grounds of availability. Only systems or applications with integration of some kind of soft computing feature or otherwise with potential for application in other intelligent earthwork systems were considered. The framework encompasses mainly but not exclusively systems with earthwork resource allocation, data acquisition or parameter estimation capabilities, which are usually the focus of these types of application. The last phase was related to the analysis and sorting of the articles into their associated earthwork phases, which allowed for the drawing of some conclusions, with emphasis on the advantages and limitation of the existent systems.

4.3. SOFT COMPUTING APPLICATIONS IN EARTHWORKS

4.3.1. Data-driven systems

A typical data-driven system consists of the application of data mining (DM) tecniques to databases. In earthworks, the aim of such systems is usually the estimation of parameters, which are unknown or uncertain during the early stages of a project. As such, most earthwork applications are based on the learning capabilities of DM tecniques on earthwork construction databases. In this context, though the type of DM technique/algorithm can be different for each case, the information contained in the database dictates the potential of the resulting system.

Successful DM applications have targeted different areas of earthwork construction, such as the one developed by Marques *et al.* (2008) and Gomes Correia et al (2012), a support system for the compaction process in constructions involving earthwork tasks. The particular feature of this system is the methodology used for the acquisition of data relatively to the compaction equipments, site materials and equipment fleet productivity indexes. In this system, the author refers to the previously mentioned GTR compaction guide to determine the productivity of the equipments under evaluation. The GTR compaction tables are subject to the DM process so as to search for patterns and tendencies in the data, with the purpose of creating a database for the determination of compaction parameters such as optimum number of passes and layer thickness. Bearing in mind that the technical guide compaction tables are a notable source of not only data concerning optimum number of passes and layer thickness depending on the type of equipment and materials being used, but also data regarding each equipment’s performance and productivity, it becomes an excellent tool for planning and controlling the quality of compaction tasks by means of procedure control.

The most relevant components of the system can be divided into two parts, both developmed under the *R* environment. The first part comprises a conventional expert system, which aims to classify compaction materials and equipment using logic rules (based mainly on “if” functions). It follows the GTR classification extensively and as closely as possible, using the same procedure a human expert would. In the case of materials, user inputs regarding the parameters obtained in the standard field and/or laboratory tests are required, as seen in the following code excerpt:

```
if (Dmax<=50 && P80>35) Classe<-“A”
else if (Dmax<=50 && P80<=35) Classe<-“B”
if (Classe==”A” || (Classe==”B” && P80>12)) cat(paste("O solo é da classe",Classe,"!\n"))
```

Where compaction equipment is concerned, the required information is mostly related to manufacturer specifications, as exemplified below:

```
Família<-readline("Qual a família do cilindro (pneus[P] / vibradores de rasto liso[V] / vibradores pés-de-carneiro[VP] / estáticos pés-de-carneiro[SP])?")
if (Família==”P”) {
      CR<-as.numeric(readline("Qual a carga por roda (kN)?”))
      if (25<=CR && CR<40) Classe<-“P1”
      else if (40<=CR && CR<=60) Classe<-“P2”
      else if (CR>60) Classe<-“P3”}
```

Considering that, especially in the case of geomaterial classification, a large number of field and laboratory tests are required, it is inferable that the system is demanding in terms of the number of user inputs. In fact, the user must provide information regarding the existent geomaterials, which is different for soil and rock materials, as well as concerning the available compaction equipment. These inputs are summarized in Table 4.1, conveying a global idea of the necessary information for achieving the GTR classification of compaction materials and equipment. In cases that include soil-rockfill mixtures, the characteristics of the latter can be similar to either soils or rocks, depending on the case. However, it is noteworthy to add that whenever there a significant percentage of fines is present in the soil-rockfill mixture, moisture control becomes essential for construction purposes.

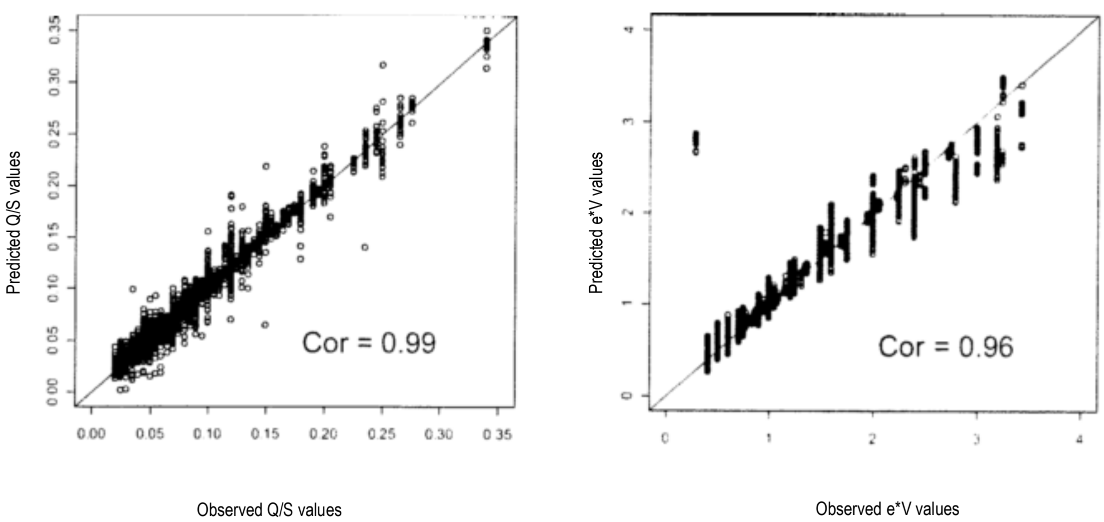
The second part of the system is responsible by its cathegorization as data-driven system. Indeed, it is comprised of the application of DM tecniques, namely artificial neural networks (ANN), to the GTR data related to compaction productivity.

Table 4.1 Required inputs for geomaterials (soil and rock) and compaction equipment

Soil	***Rock***	***Equipment***
- Maximum soil grain size (D_{max}, mm); - P80 and P2, referring to the % of material passing through the correspondent sieve; - Methylene blue absorption value measured on 0-50mm fraction (VBS, grams methyl blue per 100g soil); - Plasticity index (I_p, %); - Sand equivalent (ES, %); - Material texture; - Ratio of material fraction 0/50 mm (%); - Natural moisture content (W_n, %); - Standard Proctor optimum moisture content (W_{opn}, %); - Consistency index (I_c); Immediate bearing index (IPI, %); - Los Angeles coefficient measured on 10-14mm fraction or 6.3-10mm if unavailable (LA, %); - Micro-deval coefficient in water measured on 10-14mm fraction or 6.3-10mm if unavailable (MDE, %); - Sand friability coefficient (FS, %).	- Nature of rock; - Los Angeles coefficient measured on 10-14mm fraction or 6.3-10mm if unavailable (LA, %); - Fragmentation coefficient (FR, %); - Degradability coefficient (DG, %); - Immediate bearing index (IPI, %); - Micro-deval coefficient in water measured on 10-14mm fraction or 6.3-10mm if unavailable (MDE, %); - Bulk unit weight of dry rock sample (ρ_d); - Natural moisture content (W_n, %); - Soluble mineral content (%).	- Compactor family (Pneumatic tyred rollers, vibratory rollers, etc.); - Load per wheel (CR, kN); - Mass per unit length of the static or vibrating drum (M1/L; kg/cm); - Theoretical empty amplitude, A0 = 1000 me/M0, in which me is the eccentric moment in mkg and M0 is the mass in kg of the vibrating part excited by the eccentric (mm).

In general terms, a series of neural networks are applied to data stemming from the GTR compaction tables, with the purpose of predicting the Q/L parameter, described in Chapter 3 (see Section 2.2.1, under Compaction), for each available compactor, as a function of the material to be compacted. The information regarding material and equipment classification determined in the first part of the system is used as the reference for which the adjusted neural networks determine equipment productivity. In other words, the first part of the system classifies materials and equipment according to the GTR guide, and the second part of the system infers on the productivity of the same equipment when handling the classified materials.

The determination of productivity is achieved by using two distinct ANNs, the first of which being used to estimate the Q/S parameter, while the second the e*V ratio. While the former is a function of material type, compactor type, and required compaction energy (just as required by the GTR compaction tables), the latter depends on material type, compactor type, compaction energy and the estimated Q/S parameter. At this point, by taking a reference value for roller speed (V) and/or desirable layer thickness (e), either specified by the user or by using the suggested GTR values, it is possible to derive the missing parameter (e or V) from the estimated value of the e*V ratio. Finally, having knowledge of the Q/S, e, and V parameters, it is easily possible to calculate the Q/L value for the each compactor-material pair. Figure 4.1 depicts the performance of the DM models, showing an excellent level of adjustment and predictive capability (Cortez et al 2008).



a.) b.)

Figure 4.1 Predicted values vs observed values for: a.) Q/S parameter; b.) e*V ratio (Cortez et al 2008)

Using similar methodologies, other relevant applications feature the estimation of equipment productivity using DM on earthwork construction databases, namely application of ANN for the estimation of excavation and transport equipment productivity rates (Shi 1999; Tam et al 2002; Schabowicz and Hoła 2008) or execution time and cost in earthwork design (Hola and Schabowicz 2010). Likewise, similar DM techniques, such as multiple regressions, have been successfully applied to the prediction of excavator cycle time (Edwards and Griffiths 2000), attempting to accurately determine the duration of excavation times in earthwork tasks. Table

4.2 compiles the DM applications in earthworks, including the associated technique and estimation targets.

Table 4.2 Data mining applications in earthworks

System	***Data mining technique***	***Estimation objective***
Edwards and Griffiths 2000	MR	Excavator cycle time
Hola and Schabowicz 2010	ANN	Execution duration and cost
Marques *et al.* 2008	ANN	Compaction productivity
Schabowicz and Hoła 2008	ANN	Excavation and transportation productivity rates
Shi 1999	ANN	Excavation and transportation productivity rates
Tam *et al.* 2002	ANN	Excavation and transportation productivity rates

4.3.2. Simulation-Optimization systems

Generally, simulation-optimization systems rely on an optimizer, associated with an evaluation function, often based on linear relations or in the form of a simulation engine. While the optimizer searches for potential solutions for a problem, the evaluation function punctuates each possible solution in order to establish a measure of preferences over decision objectives (AbouRizk and Hajjar 1998).

Within this framework, Marzouk and Moselhi (2002b) presented an application focused on the optimization of earthwork tasks, associating a previously developed simulation system (Marzouk and Moselhi 2002a) to an optimization method, namely a genetic algorithm (GA). The simulation engine (ESMP) is comprised of a simulation module, an equipment database and an equipment cost application. The simulation module models the dynamic environment inherent to the equipment and resources used in earthwork operations. The equipment database provides the simulation module with the specifications of available equipment, while the equipment cost application calculates the operational time and costs of each piece of equipment, according to user inputs. Obviously, the user is required to specify the equipment available for the several operations undertaken in the project, as well as the type of operation that each type of equipment is able to perform. With this data, the system solicitates the user to choose from a series of possible optimization scenarios:

- Minimize total project cost;
- Minimize the project duration;
- Minimize the idle time of one or more specific pieces of equipment.

Since the system does not present itself capable of multi-objective modelling, should the user has the need to simulate more than one optimization scenario, the system presents a solution for each one individually. The main purpose of the system is the selection of the best possible fleet of equipment to complete a series of earthwork tasks. In order to achieve this, the GA outputs several different potential solutions (individuals), while overseeing the simulation engine. Each of these individuals is punctuated in terms of its fitness (according to the chosen objective) via simulation. This interaction between the GA and the simulation engine comprises a typical simulation-optimization system, and is illustrated in Figure 4.2.

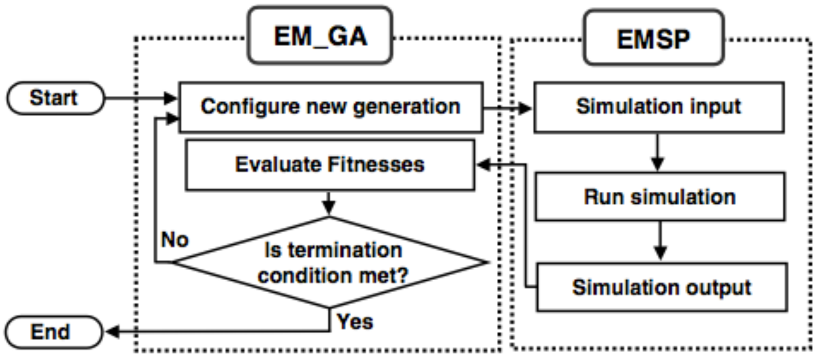


Figure 4.2 Genetic algorithm (left) and interaction with the fitness function (right) (Marzouk and Moselhi 2002b)

More recently, this work potentiated the development of a robust model for planning and control of earthwork operations has been developed (Moselhi and Alshibani 2007; Moselhi and Alshibani 2009). The system is mainly focused on continuous optimization of the equipment fleet in a construction involving earthwork operations, supported by both GAs and spatial technologies, such as global positioning system (GPS). Although still based on single-objective optimization, it includes the option of optimizing a cost plus duration parameter, attempting to minimize both simultaneously. However, unlike other cases, the particular feature of this system is the fact that it is mainly focused on continuous optimization of the equipment fleet during construction phase. In fact, it is expected to not only optimize the available resources during planning phase, but also to continue the optimization process throughout the construction phase itself, updating itself in function of the data collected by GPS, while at the same time managing every task and equipment in real time.

The GPS is used as a real time equipment control tool, as well as support for spatial data acquisition while simultaneously functioning as a spatial data analysis tool. For instance, installing a GPS emitter in transportation equipment will result in the knowledge of the positioning of that piece of equipment over time. With this knowledge, it is possible to determine the amount of hauling cycles it performs per hour, thus inferring on its on-the-job productivity. The productivity that had been estimated in design phase is replaced with the new productivity values, inferred from the continuous GPS control, calling for a re-optimization of the previously chosen equipment fleet. Since this is done in real time, the resulting system presents a superior ability to adapt to unforeseen delays during construction. The updating and re-optimization is carried automatically whenever the performances indexes of the project are lower than the ones in planning phase. Figure 4.3 depicts the system main components and their interaction.

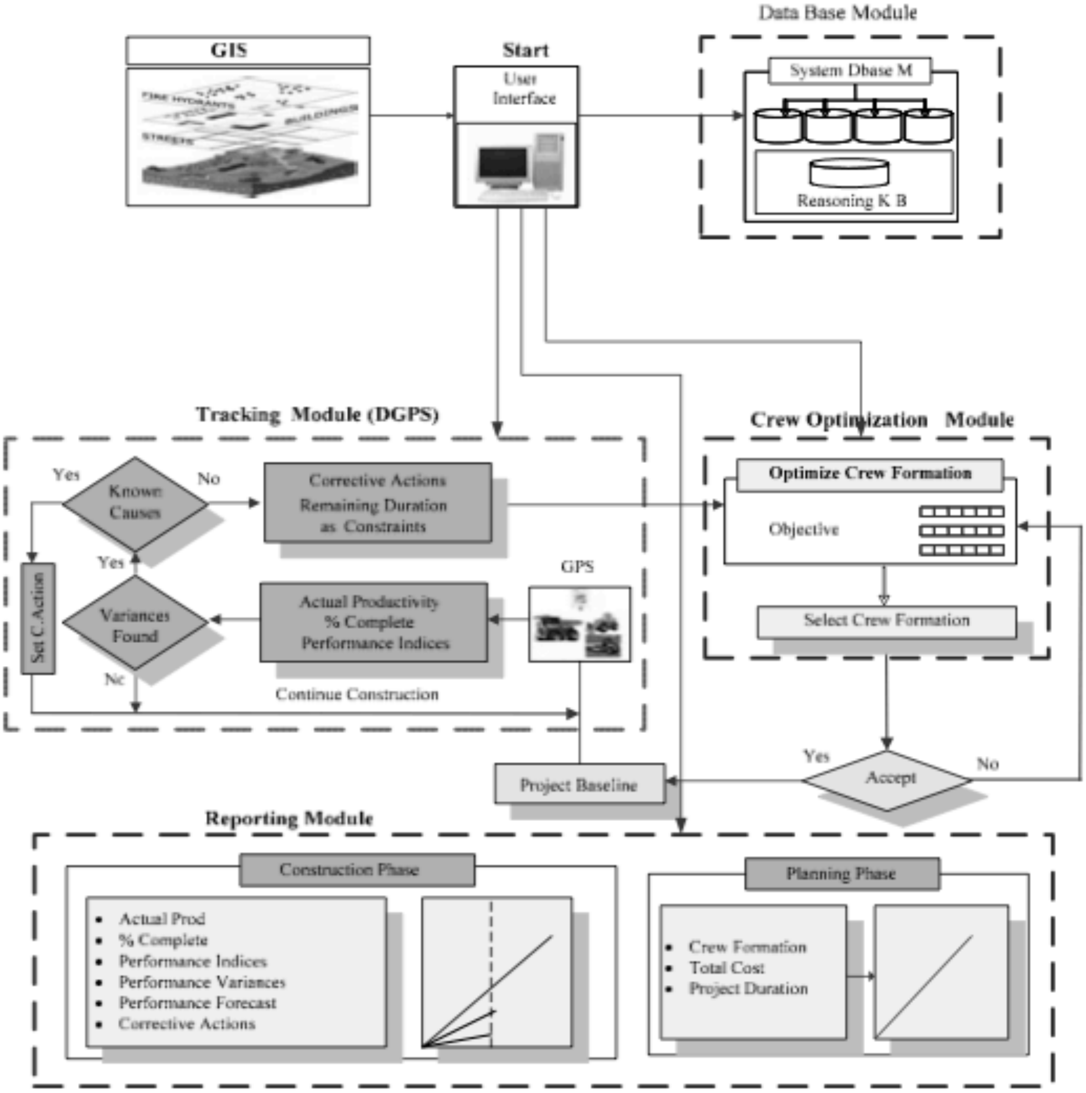


Figure 4.3 System components and their interaction (Moselhi and Alshibani 2007)

Other systems follow the same simulation-optimization principles using different forms of simulation and/or optimization methods, though most focus on specific tasks the earthwork process, namely excavation and hauling. Cheng *et al.* (2005) proposed a similar method, based on an enhanced GA optimization, while using CYCLONE (CYCLic Operation NEtwork – a broadly used commercial simulation engine) as a simulation evaluation function to assess the performance potential combinations of excavation and hauling equipment. Zhang (2008) presented a multi-objective simulation-optimization system using particle swarm optimization (PSO) for optimization and object-oriented simulation as an evaluation function. Cheng *et al.* (2010; 2011) and Luo *et al.* (2008) illustrate how Petri net (P/T net) models can be used for the management of equipment allocation in a specific earthmoving operation case. The P/T net was used to represent the dynamic constraint relationships among the various types of equipment and their functions, in order to describe the process and equipment workflow in earthmoving operations. Also focusing on transportation and hauling tasks and machinery allocation, Xu *et al.* (2011) coupled an optimization heuristic algorithm based on evolutionary strategy with a mathematical model for simulation. In order to facilitate a better understanding of the most used optimization and evaluation methods, Table 4.3 associates these to the referred systems.

Table 4.3 Applications for earthwork equipment allocation

System	*Optimization method*	*Evaluation method*
Cheng *et al.* 2005	GA	Simulation (commercial software – CYCLONE)
Cheng *et al.* 2010, 2011	P/T net	P/T net
Moselhi *et al.* 2002, 2007	GA	Object-oriented simulation
Luo *et al.* 2008	P/T net	P/T net
Xu *et al.* 2011	Heuristic algorithm	Mathematical modelling
Zhang 2008	PSO	Object-oriented simulation

Several other developments focus on partial simulation and/or optimization of individual aspects of earthwork tasks, such as excavator cycle time using fuzzy logic algorithms (FLA) (Yang et al 2003) or optimization of the least costs cut-fill route and mass haul diagrams using AI (Askew et al 2002; Mawdesley et al 2004) and swarm intelligence algorithms (Kataria et al 2005; Miao et al 2011; Nassar and Hosny 2012). Moreover, a substantial number of authors focus on the application of modern heuristics on highway alignment in early design phases (Mandow and Pérez-de-la-Cruz 2004; Kim et al 2005; Göktepe et al 2008; Miao et al 2009; Jha 2012). While not directly related to earthworks optimization, both the used optimization methods and the fact that the minimization of earthwork volumes is an essential aspect in this kind of systems demonstrate the relevance of metaheuristics on the earthwork construction context.

4.4. SYSTEM ARCHITECTURES AND APPLICATION

Early attempts of integrating AI and optimization were based on the expert knowledge acquired from planning engineers and construction equipment specialists, attempting the development of expert systems for the selection of earthmoving equipment (Alkass and Harris 1988; Alkass and Harris 1991; Naoum and Haidar 2001), optimize cut-fill routes and mass-haul diagrams (Mawdesley et al 1988) or compile information on quality control and management (Kotdwala and Basheer 1994). The knowledge in these systems is stored in a relatively simple way that allows its use in a wide variety of decision support environments. However, as traditional expert systems, they are limited to the structured rules with which they are developed. In this point of view, even though these can be used as a decision support tools in simple cases, they do not taking advantage of the full capabilities of the most recent SC techniques, such as DM and metaheuristics.

4.4.1. Architectures and used techniques

In order to structure the information in this work, a literature map built according to the research methodology described in Section 4.2 and regarding the systems analysed in Section 4.3 is presented in Table 4.4. The literature map attempts to contextualize the system and applications references in their correspondent area. It also clearly illustrates the considerably higher number of applications falling under the metaheuristics system category for excavation and transportation tasks. As previously referred, this is very likely related to the high potential of integrating simulation and metaheuristics to model real world problems, as well as the ease with which the outcome can be interpreted, validated and applied on real construction projects.

Table 4.4 Intelligent earthwork systems literature map

Technology	***Highway alignment***	***Material management***	***Excavation and transportation***	***Compaction***
Data Mining			- (Edwards and Griffiths 2000) - (Schabowicz and Hoła 2008; Hola and Schabowicz 2010) - (Shi 1999) - (Tam et al 2002)	- (Marques et al 2008; Gomes Correia et al 2012)
Metaheuristics	- (Jha 2012) - (Kim et al 2005) - (Mandow and Pérez-de-la-Cruz 2004) - (Miao et al 2009)	- (Göktepe et al 2008) - (Kataria et al 2005) - (Mawdesley et al 2004) - (Miao et al 2011) - (Nassar and Hosny 2012)	- (Askew et al 2002) - (Cheng et al 2010; Cheng et al 2011) - (Cheng et al 2005) - (Luo et al 2008) - (Marzouk and Moselhi 2002a; Marzouk and Moselhi 2002b) - (Xu et al 2011) - (Yang et al 2003) - (Zhang 2008)	
			- (Moselhi and Alshibani 2007; Moselhi and Alshibani 2009)	

In general terms, DM applications in earthwork constructions are based on the learning capabilities of AI algorithms. In fact, the feature of learning from past data and predicting its behaviour in altered or future situations has great potential for engineering applications, especially considering that it essentially simulates the process of gaining experience by an engineer, which is then used as a basis in new construction projects. Thus, DM earthwork systems rely on the existence of databases to which the learning algorithms are applied, while their outcome is limited to the type of present data and the experience gained. Nonetheless, they are susceptible of being integrated into more complex systems, as explored, even if only theoretically, by Michalewicz *et al.* (2007). These authors suggested a framework for a DM system in which a "prediction module" combined with an optimization method is capable of performing data extraction and analysis in order to determine and select the best solutions for a certain problem. The proposed system integrates the ability to be fed new data and immediately adapt and "learn" from it in real time. This fact inherently gifts the system with the aptitude for working in dynamic, every-changing environments. Even though no practical applications have been developed in the context of Civil Engineering so far, the idea of coupling DM with optimization is discussed further on. As for the examined DM systems, the fact that they focus on specific earthwork tasks, such as excavation (Edwards and Griffiths 2000) and hauling (Shi 1999; Tam et al 2002; Schabowicz and Hoła 2008; Hola and Schabowicz 2010) or the compaction process (Marques et al 2008), can be seen as a limitation. Considering the iterative nature of DM techniques, ANN seem to be the most effective one in estimating earthwork parameters, on account of leading to the best model adjustments regarding this problem. However, as previously referred, ANN systems show some difficulties when dealing with large amounts of irrelevant data, which can present a limitation when dealing with real construction data. Still, in the light of the previous discussion, its potential to be integrated in a more complex and comprehensive system is noteworthy.

Simulation-optimization systems are different from DM systems taking into account that the use of AI is not in the form of machine learning algorithms, but rather metaheuristics algorithms. Tackling the complex and dynamic environment inherent to earthwork constructions, simulation is the most common evaluation method for its capabilities and ease of interpretation on expressing real construction processes in ever-changing environments. Earthwork simulation-optimization applications can be divided into global resource allocation systems and task-specific optimization systems, depending on their optimization objectives. On the one hand, typical resource allocation systems (Marzouk and Moselhi 2002a; Marzouk and Moselhi 2002b; Cheng et al 2005; Zhang 2008) focus on the global optimization of available earthwork equipment and machinery during design phase, in function of predefined optimization objectives (i.e., time/cost minimization). A noteworthy exception for these possibilities is the system proposed by Moselhi and Alshibani (2007; 2009) which focuses on the optimization of available resources mostly during construction phase, incorporating the use of GPS to help estimate the real productivity of each available equipment and automatically re-allocating resources if these productivity rates suffer any alteration. On the other hand, task-specific optimization systems (Askew et al 2002; Yang et al 2003; Kataria et al 2005; Cheng et al 2010; Miao et al 2011; Xu et al 2011; Cheng et al 2011; Nassar and Hosny 2012) generally focus on improving the processes that form specific earthwork tasks themselves, overlooking the advantages of global optimization. These systems also fall under the same architecture of associating optimizers with simulation as an evaluation function. However, the fundamental difference between these and the previous systems lies not only on the objective of the optimization, but also on the type of simulation used for evaluating solutions. In fact, since these systems usually require a much more detailed simulation of internal processes and constraints within a specific earthwork task, the most used simulation methods are based on queuing theory and dependency relations between processes. It seems evident that GA, for their ease of implementation and intuitive interpretation, are the main choice for optimizer in earthwork resource allocation (Marzouk and Moselhi 2002a; Marzouk and Moselhi 2002b; Cheng et al 2005; Kim et al 2005; Moselhi and Alshibani 2007; Moselhi and Alshibani 2009; Xu et al 2011). Given the nature of SI algorithms, especially ACO, these seem to be more suited for route and mass-haul optimization, as is corroborated by the studied applications (Kataria et al 2005; Miao et al 2009; Miao et al 2011; Nassar and Hosny 2012). Concurrently, P/T nets and FLA are mainly limited to simulation and modelling methods (Yang et al 2003; Luo et al 2008), although there are some attempts at partial optimization of earthwork tasks, namely excavation and transportation, and tasks scheduling optimization (Cheng et al 2010; Cheng et al 2011).

4.4.2. Application in earthwork phases

In order to further this analysis, Table 4.5 and Figure 4.4 represent a system categorization map, in which the most relevant system architectures are classified into groups (architecture types) in function of their conceptual features, namely both the technologies that support them and their correspondent application areas.

This categorization identifies the different phases of an earthwork project, attempting to emphasize the absence of integration between all areas regarding the existent systems. One can easily infer that the majority of existent applications not only fall under the simulation optimization category, but also most of these focus on optimizing resources or other aspects during planning and design phase of earthwork projects (Marzouk and Moselhi 2002a; Marzouk and Moselhi 2002b; Yang et al 2003; Cheng et al 2005; Kataria et al 2005; Kim et al 2005; Göktepe et al 2008; Luo et al 2008; Zhang 2008; Miao et al 2009; Cheng et al 2010; Miao et al 2011; Xu et al 2011; Cheng et al 2011; Nassar and Hosny 2012). Additionally, a considerable percentage of the latter focus on the optimization of single aspects or tasks of earthwork projects, as opposed to attempting to integrate the optimization of all phases of the earthworks process.

Table 4.5 Matrix of application areas of existent intelligent earthwork system types

Underlying Technology		*Data acquisition and parameter estimation (Pre-design phase)*	*Planning & Design phase*	*Monitoring & Control phase*
Data driven systems		**Type 1** (Marques et al 2008); (Edwards and Griffiths 2000); (Schabowicz and Hoła 2008; Hola and Schabowicz 2010); (Shi 1999); (Tam et al 2002)		
Simulation optimization systems	*GA*		**Type 2** (Cheng et al 2005); (Kim et al 2005); (Marzouk and Moselhi 2002a; Marzouk and Moselhi 2002b); (Xu et al 2011)	**Type 5** (Moselhi and Alshibani 2007; Moselhi and Alshibani 2009)
	SI		**Type 3** (Kataria et al 2005); (Miao et al 2011); (Miao et al 2009); (Nassar and Hosny 2012); (Zhang 2008)	
	FLA and P/T nets		**Type 4** (Cheng et al 2010; Cheng et al 2011); (Göktepe et al 2008); (Luo et al 2008); (Yang et al 2003)	

Considering the different capabilities of each technology, an ideal system for earthwork control and optimization should be able to integrate all the modules in order to work throughout the whole design and construction process. However, as stated above, the systems developed so far for earthwork construction applications are generally limited or focused to only one of these areas. On the one hand, as described in Table 4.5, most simulation optimization systems focus on optimization during design phase, with the exception of the one developed by Moselhi and Alshibani (2007; 2009), which centres its capabilities in construction phase. In fact, the inability of these systems to adapt to the frequent unforeseen events associated with *in-situ* construction can be seen as a significant limitation, since most of these events are impossible to predict during design phase. On the other hand, systems based on DM techniques mostly focus on data acquisition and analysis methodologies, which in few cases are applied to design phase as a resort for estimating unknown material and equipment characteristics or parameters.

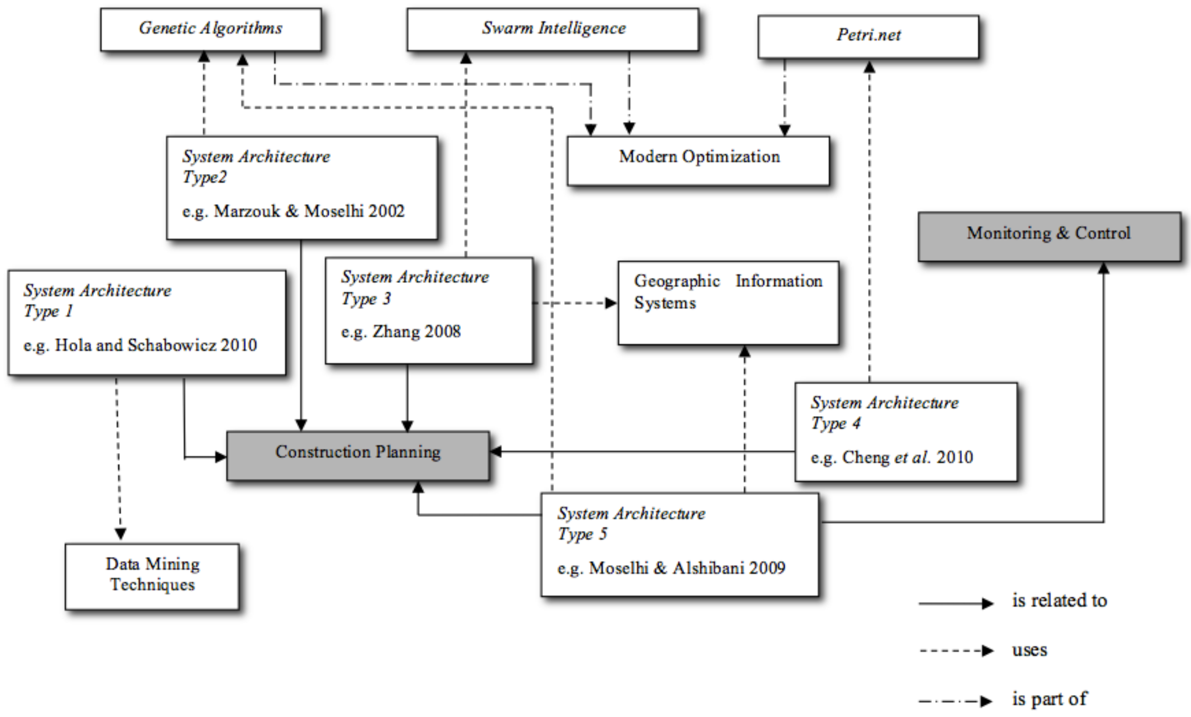


Figure 4.4 System concept map

4.5. FINAL REMARKS

In addition to the existent systems, some efforts have been made in order to predict future tendencies and possibilities for this type of systems. Kim and Russel (2003) explore the theoretical possibility for a fully automated and autonomous earthwork construction system, in which human intervention is minimal or even inexistent. The autonomous capabilities of the system include not only the automatic planning of routes for transportation equipment, but also the synchronization and interactivity between all equipment elements (from excavation to compaction tasks). However, the authors conclude that the available technology for such system still lack development and, even though some technologies that could be adapted in order to implement it are pointed out, more steps regarding technology advancements are needed in order to make this a possibility. The idea of fully automated construction is not completely new, as made obvious by the strategic agenda of the European Construction Technology Platform (2005) for underground construction. In fact, one of the long-term main objectives, among others with a deadline set to year 2030, for improving safety in underground construction is the development of totally automated remotely controlled tunnel construction equipment, techniques and processes. This has the purpose of effectively achieving the goals related to being able to carry out underground constructions both with no workers inside tunnels, reflecting the 100% automated construction goal, and with zero accidents during construction.

In order to come closer to those objectives, the next step in terms of a viable intelligent earthwork optimization system should consist of the integration of the currently available technologies (i.e., DM, metaheuristics and GIS), so as to develop a robust and reliable system capable of optimization throughout all phases of a construction project. The framework for the intelligent earthwork optimization system is presented on Chapter 5, based on the previously discussed technologies and with the potential to integrate all the mentioned areas. This includes knowledge extraction from databases as a mean to support an optimization system capable of not only planning and optimizing earthwork construction tasks, but also be flexible enough to allow the updating of site and equipment conditions during construction phase, re-optimizing the system should the efficiency of the earthworks process fall short of what was estimated in design phase.

Chapter 5

DEVELOPMENT OF THE EARTHWORK OPTIMIZATION SYSTEM

5.1. INTRODUCTION

Ideally, earthworks should be optimized automatically. Considering the nontrivial characteristics of earthworks optimization (e.g., large search space and conflicting goals), conventional Operational Research (e.g., linear programming) and blind search methods are infeasible. As an alternative, metaheuristics are an interesting solution within this domain, since they are capable of searching interesting search space regions under a reasonable use of computational resources. Indeed, several studies have adopted metaheuristics to this domain, such as genetic algorithms (Marzouk and Moselhi 2002b; Moselhi and Alshibani 2007; Xu et al 2011) and swarm intelligence (Kataria et al 2005; Zhang 2008; Miao et al 2011; Nassar and Hosny 2012). However, many of these applications focus on single tasks or partial processes that comprise earthworks, i.e., excavation and hauling (Kataria et al 2005; Xu et al 2011), in an attempt to deal with the high complexity of the problem. For this reason, 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 and Moselhi 2002b) 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). However, as one can easily infer in nontrivial multi-criteria optimization problems, 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. 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. And additional criteria could be used to support the final decision, such as environmental aspects, which can be assessed by the determination of carbon dioxide emissions in each solution.

5.2. SYSTEM ARCHITECTURE

The proposed system architecture consists of three integrated main modules with capabilities to acquire and manipulate data from each phase of an earthwork project. Figure 5.1 depicts the module organization and information flow, while Table 5.1 summarizes the technologies, tools and purposes associated with each of the three modules:

- Equipment Module
- Spatial Module; and
- Optimization Module.

Table 5.1 Modules, technologies, tools and functions

Module	***Technology***	***Tools***	***Function***
Equipment	Data Mining	R, Rminer	• user inputs; • estimation of productivity & costs
Spatial	Geographic Information Systems	R, GIS, GPS	• modelling of construction site; • path finder
Optimization	Metaheuristics	R	• (near) optimal selection of equipment fleet depending on availability; • (near) optimal equipment fleet allocation throughout construction phase; • return output to user.

In general terms, the Equipment Module is responsible for receiving the user input for available equipment/plants, while calculating or retrieving equipment costs. Simultaneously, this module should have a close interaction with the data mining (DM) models for determining the productivity rates for available rollers. This model is built upon the data comprised in the *Guide des Terrassements Routiers* (GTR) compaction guide, such as material characteristics, roller specifications (i.e., weight per drum length and maximum amplitude, in the case of vibratory rollers) and construction specifications (i.e., required compaction energy or specific layer depth). Given this input, the model is then capable of classifying both the material and roller types, as well as retrieve information regarding number of compaction passes and maximum productivity. Moreover, given the availability of further construction data, the DM models in this module can be expanded to include the whole construction equipment, such as spreading, transportation and excavation equipment.

On the other hand, the Spatial Module allows the user to input a functional model of the work site by user input using a geographic information system (GIS), namely including all the possible work fronts and potential equipment trajectories/paths. Simultaneously, the path finder algorithms integrated in most GIS software can be used to determine the best routes or trajectories for transportation equipment, by making of both the created work site model and the equipment specifications from the equipment module. This results in the optimization of the workflow in the construction site. Furthermore, depending on the availability of global positioning system (GPS) equipment, by including GPS receivers in the active earthwork equipment during construction phase and associating these with the GIS software, it becomes possible to determine the actual equipment work rates. This allows the system to automatically update and optimally adjust itself in real-time as the construction process goes on. For instance, consider the possibility that the real rate of transport equipment (such as a dumper truck) is inferior to what was originally predicted in the modelling phase in terms of travels between excavations or borrow sites and work fronts per hour. By monitoring and re-evaluating the real rate of this equipment via GPS (i.e., number of trips per hour), the system would then be able to perform adjustments, such as relocating another piece of equipment, in order to keep the original work flow. Depending on the availability of number of GPS equipment present on the site, this function is ideally susceptible to be extended to all the working equipment in the fleet.

Finally, the Optimization Module works by means of an optimization algorithm (i.e., a GA), which attempts to find a near optimal solution for the problem of determining the best possible equipment fleet and its optimal distribution throughout the work area. The optimization is carried out bearing in mind both construction time and costs, which are often conflicting interests in earthwork tasks, defining the problem as multi-objective optimization with conflicting objectives. Additionally, considering that the optimum equipment locations are not static over time, since as the tasks in one work front are completed the equipment should be optimally reassigned to other work fronts, one could further define the problem as a dynamic multi-objective optimization with conflicting objectives. As such, both the Equipment Module and the Spatial Module feed the Optimization Module with the necessary data to carry out the search for the near optimal solutions for the problem. The latter is able to evaluate the performance of each solution by means of a fitness function regarding both costs and time. The simulation method may be, for instance, based on object-oriented simulation of the whole construction process for each potential solution, allowing for the determination of both costs and construction time in each equipment fleet configuration. After the best solutions have been considered and evaluated, the Optimization Module presents the user with the best-found solutions through the User Interface as the output for the system.

It is important to refer that, in what concerns the optimization and performance evaluation processes, the main parameters can be seen as being directly related to the total roller rate that is being considered for each work front (in each iteration). In fact, if the earthwork process as a whole is looked at in a production line point of view, in which the work rate of the whole production line corresponds to the work rate of the last process (in this case, compaction), which, in turn, depends on the rates of all precedent processes, then maximizing production in an earthwork production line means maximizing the productivity of compaction throughout the whole construction. In this context, the optimization of the earthwork process must take into account not only the optimal allocation of compaction equipment, but also the determination of a compatible production line preceding the compaction task. As such, the work rates of spreading, transportation and excavation equipment (including excavation from borrow sites), as well as equipment used in any intermediate tasks (i.e., soil treatment equipment, water trucks), should be determined taking into account the roller rate in each work front so as to allow for the constant work flow in those areas. In other words, each work front as to be “fed” just enough material for the rollers to constantly work at their maximum rate, minimizing their idle time, thus demanding the synchronization of the whole equipment fleet, including transportation and excavation equipment.

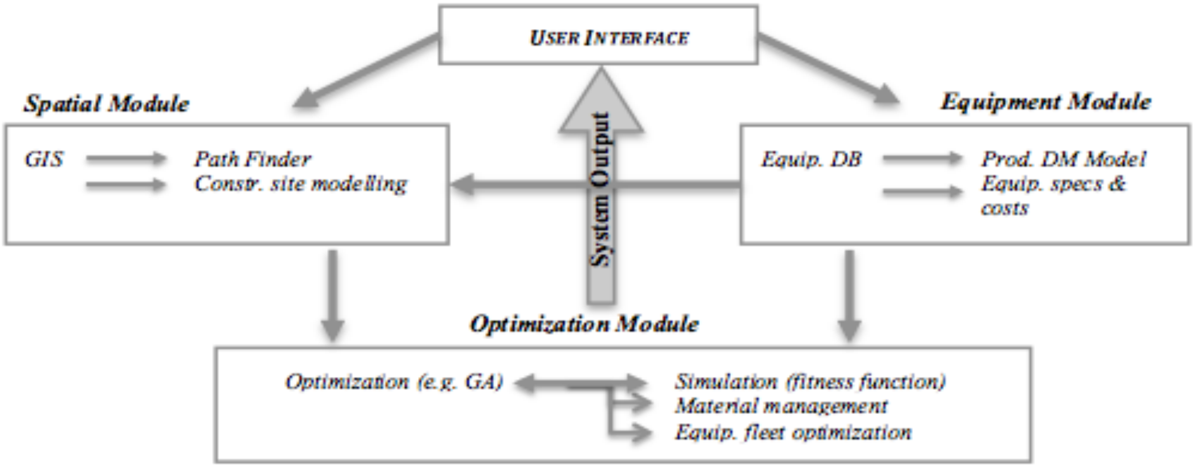


Figure 5.1 System architecture

5.3. SYSTEM MODULES

5.3.1. Equipment Module

The equipment module is based on the system previously developed by Marques, Gomes Correia and Cortez (2007; 2008), which has been described in detail in Chapter 4 (see Section 4.3.1). Using the data from the GTR compaction tables, in conjunction with user inputs regarding the compaction equipment (e.g., available rollers and specifications or GTR classification) and embankment fronts (e.g., material type, preferred layer thickness), this module is able to assess the productivity of any roller type in each embankment front. The original system includes the classification of materials and rollers depending on their specifications in the form of an expert system. However, in order to facilitate the system's input requirements, as well as simplify development of the intelligent optimization system at this phase, the inclusion of the automatic material and equipment classification feature of the original system was assigned to future work. Since this notion was already taken into consideration since early stages of development of the equipment module, the necessary steps have already been taken so as to allow for an easy addition of this feature. Be that as it may, since this feature forces the user to insert all the required data into matrixes prior to running the optimization system, thus increasing input requirements, it should be included as an optional input method for the user, rather than mandatory. Nonetheless, this can be facilitated by the creation on a graphical user interface (GUI), as discussed in section 5.3.4.

As such, in the present system, the classification of materials and rollers is left for the user, who should input the final class of each material and roller (according to the GTR guide) to the system. The system should then be able to automatically retrieve the productivity for each combination of roller-material in each embankment front, depending on other specifications such as layer thickness. However, this module does not have to be limited to the estimation of compaction parameters, as it can be extended to estimate productivity of more equipment types, in cases where user input is missing. This requires a robust database, either built from real construction data or originated in construction guides or manufacturer specifications. In this sense, some developments were carried out during this project, which show the viability of this concept, assuming a complete and reliable construction/manufacturer database is available. The following sections present a study based on the application of artificial neural networks (ANN) in an earthwork construction database. The DM algorithms have been tested in order to achieve the best possible adjustment to the data. Results show not only the potential of such applications, but also weight of each earthwork construction process and its influence on the production line, as well as on the final construction rate. Consequently, these also emphasize the importance of optimization in terms of resource distribution in the production line of each work front.

Database

This DM application on earthwork data is founded on the learning and predictive capabilities of AI algorithms. However, it is noteworthy to point out that this type of application relies on the existence of databases to which the learning algorithms are applied, meaning that their outcome is also limited to the type and quality of present data, as well as the gained knowledge/experience. Thus, the availability of proper data becomes essential for successfully building, training and testing of DM algorithms in earthwork construction.

In this work, a subset from a database devised from the earthworks of a Portuguese road construction site was used for that purpose. The original database includes the description of several years of earthworks construction, broke down into the daily activities of the available mechanical equipment. In this application, the data subset regards the activities of earthwork equipment throughout 6 months of construction phase, featuring around 1250 entries (after data preparation) with information on date, work hours, atmospheric conditions, number and distance of load trips and resource types, as depicted in Table 5.2.

Table 5.2 Values extracted from the available earthwork construction database

Date	Work Hrs.	Atm. Cond.	Nr. of Loads	Excav. #	Load Zone	Unload Zone	Resource Type	Transp. Volume
9/2/11	7	Rain			7+850	8+625	Excavator50T	
9/2/11	3	Rain			7+850	8+625	Roller15T	
9/5/11	9	Sun	37	20/871	13+750	12+250	Dumper40T	481
9/5/11	9	Sun	39	20/871	13+750	12+250	Dumper50T	634
9/5/11	9	Sun			13+750	12+250	Tractor40T	

Application results

As previously stated, a DM model is dependent on the available data. As such, in order to achieve an ideal model with predictive capabilities for a specific target variable, it should be built using all the variables with some degree of influence on the value of the target variable.

However, data regarding those variables is not complete or even available in many cases, especially when dealing with real world data originated in earthwork constructions. Moreover, an exceedingly high number of variables will generate too much complexity regarding the search of relations and patterns amongst variables, often lowering the predictive capability of the model. In order to surpass these issues, some non-conventional methodologies were used in this application, as discussed further ahead.

In this work, one of the main purposes for the application of DM algorithms to earthwork data was the creation of a model with predictive capabilities regarding the final construction rate of the whole process. Considering the traditional earthwork construction sequence, the final construction rate corresponds to the rate of the production line itself, which coincides with the work rate of its last task. As such, the emphasis is given to the final rate of compaction equipment, corresponding to the last task of the earthworks process, which is determined while taking the excavation, transport and spreading equipment rates into consideration. Yet, as a consequence of the high number of variables with influence on each process comprising the earthwork production line, building a single DM model targeting the rate of compaction equipment would neither be efficient nor effective, since it would not display a suitable predictive ability. Instead, two sequential models were developed, the first one targeting the prediction of the daily number of load operations using excavation and transportation plants, which was then used as an input for the second model regarding the rate of spreading and compaction plants. Note that, since the prediction of the number of loads is based on real construction data, it already takes into account the durations associated with loading, hauling and return trip. The same occurs with the prediction of final compaction rate, since every associated task, such as spreading or controlling the layer water content, is inherently being taken into account. Figure 5.2 shows the obtained model results for the target variables, as well as the training variables used to predict them and their relative importance.

Model assessment was mainly based on the value of the error defining the degree of learning of a given model, as well as the correlation between the observed and the predicted values (Hastie et al 2009). The used metrics were root mean squared error (RMSE) and correlation coefficient (R^2) (See Chapter 3). Results were obtained using the rminer package (see Chapter 3) for the R tool (Cortez 2010b). The developed models feature RMSE and R^2 values equal to 8.325 and 0.855 for the first model (number of loads by transportation equipment), and 26.377 and 0.980 for the second model (compaction rate), respectively. Moreover, none of the models showed a mean absolute deviation above 12%. These values were deemed adequate seeing as the data originates from a real construction environment.

The results obtained from these models demonstrate the importance of equipment allocation in earthworks. In fact, the analysis of Figure 5.2b shows that the number of loads from transportation equipment as predicted by the corresponding model is the main factor influencing the prediction of final compaction rate. In other words, the daily work rate of the production line preceding the compaction process is the main factor affecting its progress, affecting the development of the whole construction.

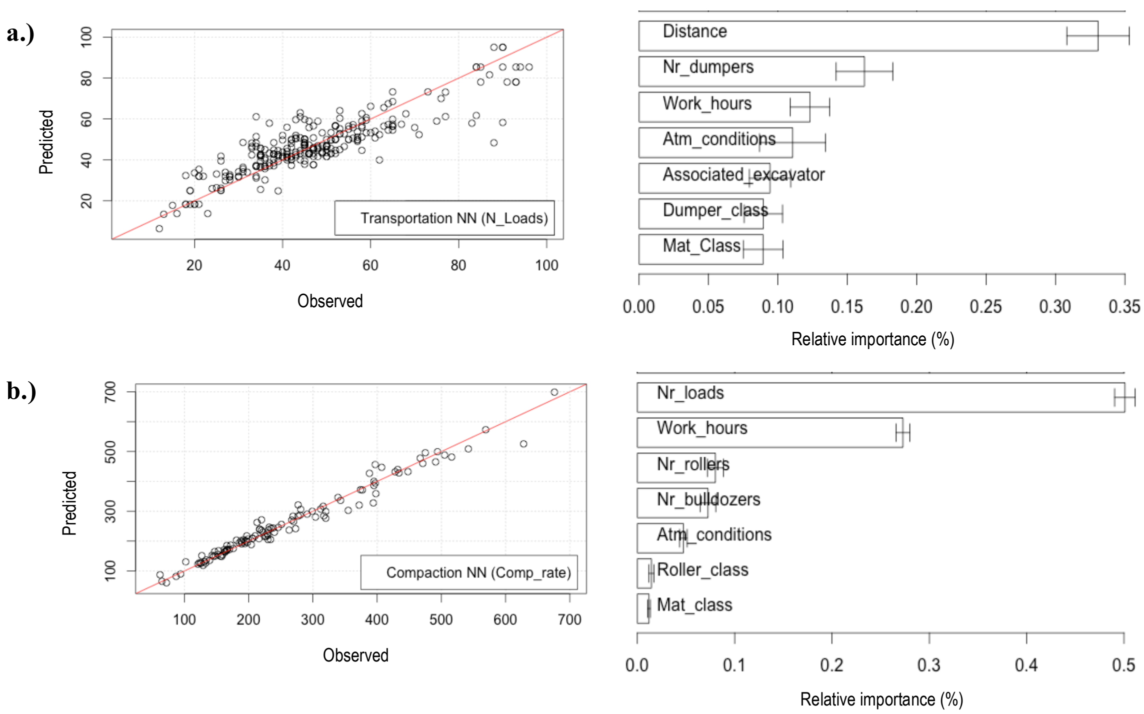


Figure 5.2 Results yielded by prediction models and the observed values (left), including relative importance (%) of variables for each model (right) for: a.) Number of transport loads, b.) Compaction rate (Parente et al 2014).

This is often verified in real constructions cases if the productivity of excavation and transportation plants does not match that of the compaction plants. On the one hand, should the productivity of excavation and transportations plants be inferior to that of compaction plants, the latter will sustain high idle equipment times, reducing the global work rate. On the other hand, a higher productivity regarding excavation and transportation will result in an overflow of material on the compaction front, which cannot be timely compacted, ultimately becoming an obstacle for the construction of the embankment.

In order to verify this occurrence, while also supporting the significance of optimal allocation of earthwork equipment, the developed sequential models were used to predict the work rate of earthwork equipment using a different construction setup. In the new setup, excavation and transportation plants were virtually and randomly reorganized throughout the same construction site, without altering the positions of excavation or compaction fronts. Compaction conditions were fixed, including compaction and spreading plants, as well as their distribution throughout the compaction fronts, so as to facilitate comparison with the original setup. Initially, the first model (transportation ANN) carried out the prediction of daily number of loads each excavation-transportation plant would carry out, which is a function of the variables shows in Figure 1a (in which the distance from excavation to compaction fronts, as well as the associated number of dumpers, represent the factors with most weight). The results from this model were then grouped regarding the target compaction front associated with each excavation-transportation plant, and inserted into the second model (compaction ANN), which estimated the average compaction rate in each compaction front. The compaction values found in the GTR compaction guide were used as a reference to control the maximum predicted compaction rates.

Results showed a decrease of approximately 15% in the final compaction rate corresponding to the new equipment distribution. The compaction values found in the GTR compaction guide were used as a reference to control the maximum predicted compaction rates. This decrease in the final compaction rate of the new production line, when compared to the original setup, is a result of mismatching work rates between excavation and compaction fronts. It is important to bear in mind that there is a limit to the maximum rate of a specific compaction plant, which is likely being exceeded in some work fronts for the current equipment distribution, while other work fronts are not receiving enough compaction material to achieve a satisfactory compaction rate. Thus, the compactors at the end of the production line are being forced to suffer idle periods, ultimately lowering the global compaction rate. Even though this exercise resulted in a decrease in the final compaction rate for the new equipment distribution when compared to the original setup, its purpose was considered fulfilled. However, only one distribution will be equivalent to the best possible work rate with the current available equipment fleet, which can be found by automatic optimization in terms of resource allocation, as discussed in Section 5.3.3. In this perspective, these results show the importance of equipment fleet optimization throughout construction phase, as well as the effectiveness of DM tools in earthwork construction projects.

5.3.2. Spatial Module

The spatial module is founded upon GIS technology. Its primary purpose is to gather the necessary information, generated from spatial data, to assist the proper optimization of transportation equipment distribution and workflow. This data is associated with the optimal routes, distances and cycle times between excavation and embankment fronts, as well as with spatial models of earthwork construction sites, which include information regarding the relative positioning between fronts and the possible routes that connect it. On the other hand, the spatial module is expected to provide a basis for spatial input and output visualization, enhancing the system with the possibility of receiving GPS data (if available) during construction phase of earthwork projects, as a means of updating the productivity estimated during project phase if necessary. Considering the dynamic and hard-to-predict nature of earthwork projects during construction, these features add the necessary versatility to the system to deal with unforeseen events or problems.

The initial objective of this project included the integration of both the spatial and the optimization modules of the system. This would not only integrate the GUI of *QGIS* into *R*, but also facilitate data sharing between the software, since some of the equipment specifications inputted in the equipment module could be imported into the spatial module, while the outputs from the latter could also be directly imported to the optimization module. However, after experimentation with the software, it was concluded that full integration between *QGIS* and the *R* environment is not feasible, due to incompatibilities between software versions. This is due to the fact that, in order to achieve a proper integration between *R* and *QGIS*, the installation of an interface package (i.e., *rpy2* interface) is needed, which would compatibilize the object-oriented programming languages of *R* with the Python language upon which *QGIS* is based. Yet, the available versions of this interface would force the installation of an earlier version of both *R* and *QGIS*, so as to allow for compatibility between the software. As such, following through with this option could hinder the proper functioning of the *R* packages used for the DM and optimization modules, as these were developed on a later version of the former.

Nonetheless, the optimization system built in *R* and the corresponding optimization module (see Section 5.3.3) feature the necessary flexibility to allow a manual insertion of the spatial data for the optimization process, including shortest routes, haul distances and cycle times, in a practical way. Effectively, the main subsequent limitation of the system, stemming from the unsuccessful integration of GIS software, is associated with the output representation (visual support) of final equipment positions over time, and overall user-friendliness associated with the GIS GUI support. In this context, an origin-destination (OD) cost matrix can be used to abridge the main input required from the GIS software in order to allow for the correct functioning of the system. This type of spatial data can be achieved by one of two ways:

- Manually determine or estimate haul route distances or travel times;
- Automatic estimation and generation of an OD cost matrix via GIS software.

The first method can be carried out when a good understanding of the construction area already exists, either as a result of the availability of quality data during design phase, or during any part of the construction phase of the project, in which there is a more detailed grasp of the duration of haul trips, as well as potential restrictions/obstructions regarding the available routes. In practice, it consists of manually creating an OD cost matrix based on observation or estimation of the average time that trucks require to go from each point of origin (excavation areas) to each destination point (compaction areas). Ultimately, an OD cost matrix compiles the information of how long an average truck requires to transport material from each excavation to each embankment front. However, this process is obviously much more demanding and subjective when compared to the second referred alternative. Furthermore, being carried out by human observation and common sense, optimal routing is not guaranteed.

Whereas the second method features some advantages, its main disadvantage regards the requirement of some basic knowledge of GIS software. In this method, one basically builds a model of the construction area using whatever information is available, either during project and construction phase. The model includes the origin areas, as well as destinations, and the indentification of all possible routes for the transportation equipment inside or outside the construction site. It then becomes possible to use the automatic *path finder* tools in GIS software to automatically generate the OD cost matrix and import it into *R*. This methodology is to be explored in detail further ahead in the present section.

Obviously, even if a model is built during design phase, where not enough information is still available, this type of input is susceptible of being updated during construction phase. The user can easily modify the matrix values at this stage by:

- adding a new point of origin (i.e., excavation or borrowing site) or destination (i.e., as a result of updating site conditions or dealing with unpredictable situations);

- updating/modifying the value of distance or travel duration from each point of origin to each destination point;
- updating information regarding maximum speed in each possible path in the GIS;
- adding new potential routes for the equipment or remove any of the already established ones (e.g., after concluding that one of the paths is no longer suitable).

In the light of the previous arguments, the process of using GIS software to obtain an OD cost matrix is described in the steps summarized below. In the example that follows, *ArcGIS* was used, since it simplifies the process of obtaining the matrix directly, without the need for any programming any additional loops or commands. Additionally, *Google Earth* is used in the process as a support for obtaining a partial aerial image of the construction site area. This image can then be used as a reference, over which all potential transportation equipment routes can be identified and characterized. The subsequent spatial analysis done by the GIS software will then assess the optimal routes for the transportation equipment for each origin and destination points, according to the available alternatives. It is to note that, this can be achieved by other methods (as previously referred), such as building graphs that represent the construction site and its conditions or manually adding the reference points and distances. Although these methods call for much more effort and often require a high simplification, as they are not perfectly suited to fully represent real world situations.

In this context, the first step to generate the necessary data for proper equipment allocation is obtaining or generating a reference surface of the area under analysis. This should include the locations of work fronts, borrowing sites and equipment storage area for a given phase of the work, as well as all potential paths and infrastructures that enable communication between these areas. In this case, it consisted of an aerial image of a construction site in Alijó, Portugal, as exemplified in (Figure 5.3), obtained from *Google Earth*. To exemplify the process, this aerial image corresponds to a partial area of the construction site. This method requires the collection of further information on the coordinates of at least two points so as to allow for the usage of the ArcGIS georeferencing tool. So that this can be achieved, one must import the image into the GIS by using the *Add Data* function, followed by setting the coordinate system used as a reference for the coordinate points. Given that *Google Earth*'s internal coordinate system (latitude/longitude) corresponds to the World Geodetic System 1984 (WGS84), this information is inserted in the *Data Frame Properties* of the project. The georeferencing of the image is accomplished by adding the known dimension control points in their corresponding location. Coordinate values (in degrees, minutes and seconds) in the Northern and Eastern quadrants are considered positive, while Western and Southern coordinates are inserted as negative values. By georeferencing an image, the real size of the area it represents is being defined, meaning that the distance from one point to another corresponds to its real value. This is essential for the creation of a network, since the latter is built upon the same georeference, thus containing the information on real distances and path lengths/widths for the mapped area.

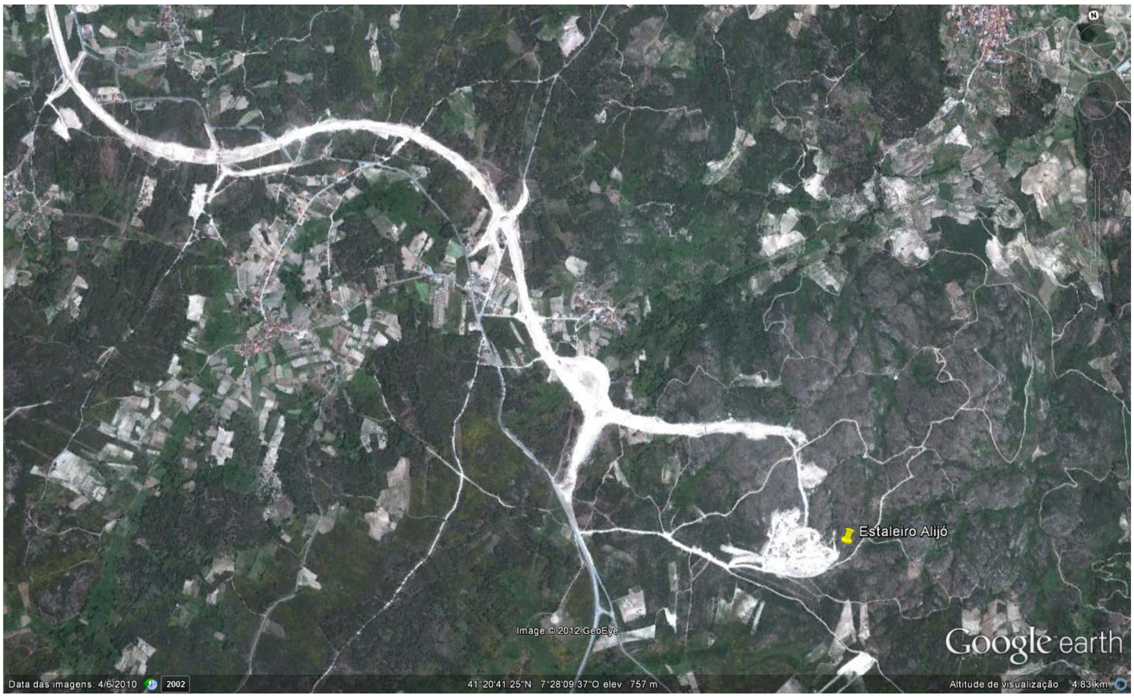


Figure 5.3 Aerial image of construction site (*Google Earth*)

The creation of a network in most GIS requires the generation of a *File Geodatabase*, which in this case was accomplished via the *ArcCatalog* tool. In turn, the *File Geodatabase* integrates a *Feature Dataset*, referenced in the same coordinate system mentioned in the previous step, which includes *feature classes*, such as lines (used to represent paths or trajectories) and points (used for defining excavation, embankment and borrowing fronts). These elements comprise the network itself and are added on a different layer over the original image, using the same georeference as the latter. As such, the image is used as a reference for the position and dimensions of roads, paths and construction zones where the mechanical equipment can work or travel. Roads and paths, associated with line feature classes are overlayed on the previously georeferenced image, corresponding to the red lines depicted in Figure 5.4. The georeferenced image also serves as a reference for indentifying work fronts in the construction site (defined by point feature class). Some of the necessary attributes to analyze subsequent networks can already be defined in this phase. Attributes correspond to columns comprised in the attribute tables of the corresponding line/point features. These can be discretized/divided into three elements:

- Distance, in km, which, since it is preset to decimal degrees (coordinate system), must be converted into metric system values, so as to simplify subsequent calculations. For conversion purposes 1 decimal degree is considered equivalent to 111 Km;
- Average speed, in km/h, which is intended to correspond to the average speed that the transport equipment can travel at, depending on type of terrain and haul path conditions of the construction site. In this example, three types of terrain were considered: paths/trajectories inside the construction site that were of limited size were attributed a maximum speed of 20 km/h; wider paths/trajectories inside the construction site were attributed a maximum speed of 30 km/h; the paths/trajectories outside the construction site (i.e: roads) were attributed a maximum speed of 50 km/h;

- Travel time, in minutes, with a corresponding value of $60\times\frac{Distance}{Average\ speed}$. The value 60 is used to convert hours to minutes.

Figure 5.4 Definition of all possible trajectories for transportation equipment (*ArcGIS*)

Although this may be considered an intermediate step, the application of the *Integrate* function (included in the Data Management Tools Toolbox – Feature Class) to the line features (trajectories) in this phase guarantees connectivity between all elements that comprise it. The importance of this step is related to the difficulties that arise in following phases, where it is often verified that the elements (polylines) comprising trajectories/paths are not connected between themselves and, consequently, are not included in the analyses for the determination of the best route for the transport equipment. As such, this will prevent having to execute this action in subsequent phases, potentially forcing the repetition of other steps, namely the creation of a new *Feature Dataset*.

Having included all the necessary elements for the network analysis process, the next stage is generating the network itself. Since all the line features are already overlayed and connected (due to the *Integrate* function), they are converted into an independent network. It should be noted that, during the generation of the network, the definition of the attribute to be used as impedance (cost) for the subsequent analysis is required. As previously mentioned, this can be either defined as travel time or distance, according to user preference, since the optimization module was developed with this option in mind. In this example, travel time was the used impedance variable. This phase also includes the appointment of restrictions, such as one-way routes, for any defined trajectories/paths in the network.

Completed the generation of the network, it is then possible to carry out the desired analysis. In order to determine which paths/trajectories are the shortest of the transport equipment between the borrowing areas and the work fronts (given the defined conditioning factors), a *Route Layer* is added using the *Network Analyst Toolbox*. This command requires the path finder algorithms, used in the GIS software, to qualitatively evaluate each of the potential trajectories as a function of the selected impedance (travel time), returning the best path between two points in the current model as output. As illustrated in Figure 5.5, by taking the points associated with an excavation and an embankment front as origin and destination, respectively, the GIS returns the best path that minimizes travel time. It is also possible at this point to add temporary or permanent barriers that may pose a hinderance or obstruction to the circulation of vehicles in the form of dots, lines, or polygons. For instance, these can be used to simulate changes in the conditions of some of the paths/trajectories along the site.

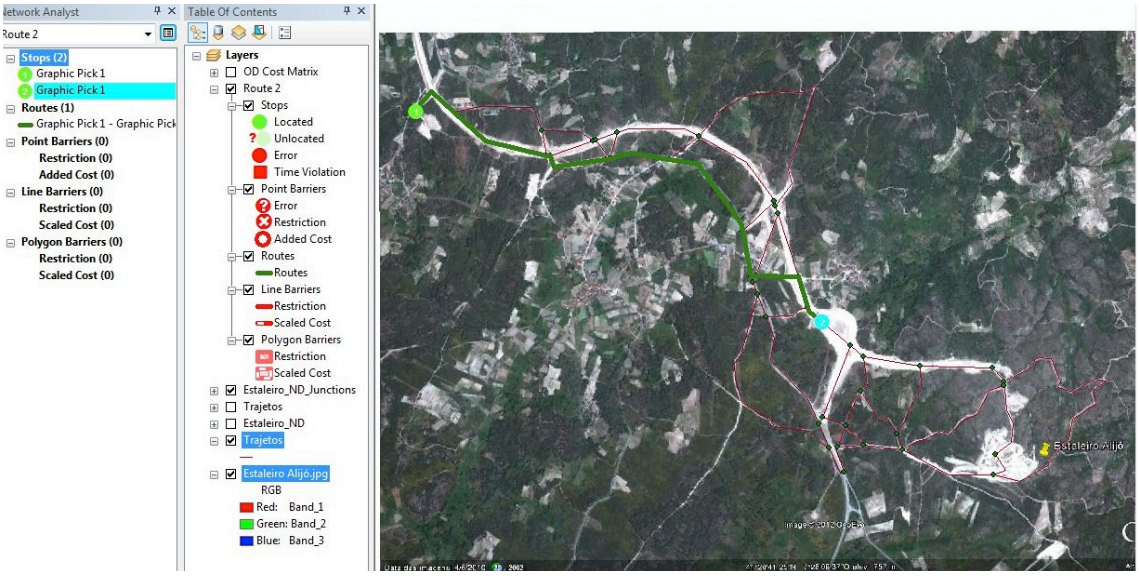


Figure 5.5 Route layer, including identification of shortest path from an excavation to an embankment front (*ArcGIS*)

However, in most cases, it is necessary to assess the travel time from one or more excavation or borrowing areas to several other work fronts. In this case, it is possible to create a cost matrix between the origin and destination points, using the selected measure as impedance (Figure 5.6). This step is accomplished by generating an *OD Cost Matrix*, using the *Network Analyst Toolbox*, by defining the origin points as the intended excavation or borrowing areas and the destination points as the embankment fronts or dump areas. Although it does not visually represent each path, *ArcGIS* uses the shortest routes as a reference, outputting the travel times (or other impedance value) for all origin and destination points in the form of a matrix. The latter can then be imported into *R*. Although ideally this action should be automatic if integration is achieved, it can also be done manually by exporting the resulting OD cost matrix to any of the formats recognized by *R* (e.g., spreadsheet format).

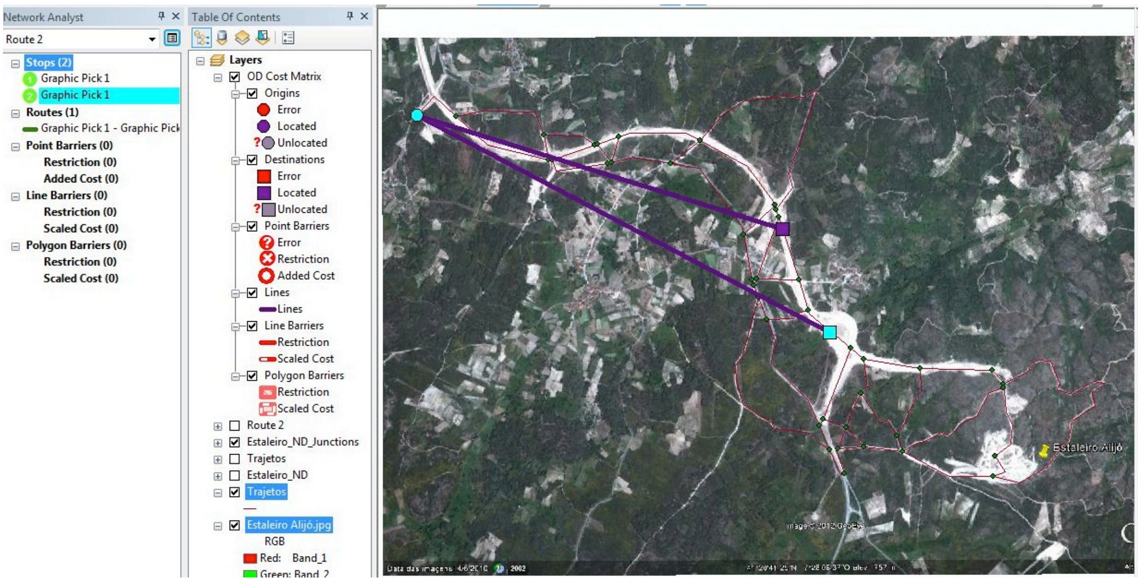


Figure 5.6 Origin-destination cost matrix with identification of one excavation front (origin) and two embankment fronts (destinations) (*ArcGIS*)

In practical terms, it is easy to infer how the information compiled in the OD cost matrix can be used in the determination of transportation equipment productivity. In fact, access to the equipment specification provides an understanding of its on-the-job capacity, which, together with the gathered GIS data regarding travel times between work fronts, makes it possible to estimate the productivity of one truck under given conditions. Moreover, as discussed further ahead in Section 5.3.3, bearing in mind that, in simple terms, the maximum productivity of each task (in an earthworks production line) should not exceed the productivity of the preceding or succeeding tasks, the optimal number of trucks in each transportation plant can be determined. Leveraging this knowledge, it then becomes possible to allocate the available transportation equipment as necessary, depending on the productivity requirements for each case and the availability of the other types of equipment, which support the transportation task, namely excavators.

5.3.3. Optimization Module

Taking into account an optimization point of view, earthwork construction can be described as a number of production lines based on resources and dependency relations between sequential tasks. The resources correspond to the mechanical equipment that is essential for the development of the project, namely excavators, dumper trucks, bulldozers and compactors, while the sequential tasks are related to the associated processes, specifically excavation, transportation, spreading and compaction, respectively. The speed at which the latter can be completed depends on the amount of the former being allocated into each task. In other words, the work rate in each sequential task (in this case often measured in volume of handled material per hour, m^3/h) can be manipulated by increasing or decreasing the amount of associated resources allocated to it. This means that earthworks are strongly susceptible to optimization, which is aimed at minimizing both execution cost and duration. The multi-criteria include conflicting properties: one can decrease execution duration by increasing the amount of allocated resources to a task, but such results in an increase of the associated execution costs and vice-versa. It should be noted that the costs related to fuel and machinery maintenance (indirect costs) are substantial, thus solutions with longer durations are not necessarily less costly.

The tasks that comprise earthwork projects have a set of specific characteristics in this context, of which the focal point is interdependency. Indeed, earthwork tasks are not only sequential, but also the work rate of each of them is always limited to the work rate of its preceding task. For instance, the dumper trucks cannot undergo the transportation of soil if the latter has yet to be excavated and loaded into them; and bulldozers cannot spread soil into layers so as to allow compaction if the material has not been brought to them by the dumper trucks, and so on. Furthermore, when dealing with sequential and interdependent tasks such as these, the speed at which a single production line can carry out its work is equivalent to the work rate associated with its last task. In this context, maximizing the work rate in the final task (in this case, compaction) would correspond to a solution with minimum execution time for a production line. However, it is noteworthy to emphasize such allocation is limited by the available equipment and also by the site conditions, such as space restrictions in excavation or compaction areas (usually designated as fronts). To fully take advantage of the available resources, one must guarantee that the allocated compaction equipment is fed enough material so as to allow for constant production. In other words, the work rate in all tasks prior to compaction (excavation, transportation and spreading) must be equal or similar to the work rate obtained in the associated compaction front. Should the work rate of a task fall short of the work rate of succeeding tasks, then the productivity of the whole production line will be limited to the one obtained in that task. This keeps the equipment from reaching its maximum potential in terms of work rate, i.e. by forcing it to idle while waiting for material. Therefore, it is essential to control the work rate in each task within a production line.

Naturally, an earthwork construction is not depicted in a single production line, but rather in several independent production lines working simultaneously. Each of these production lines is associated with a compaction front, since that is the final stage for handling the geomaterials. Moreover, there is one more characteristic specific to these production lines that significantly increases its complexity. As construction ensues in several simultaneous production lines, compaction work will come to completion in one production line at a time. At the point when one production line has completed its assignment, the associated equipment is no longer contributing towards the completion of the earthwork project, thus calling for its reallocation into either an existent or a new production line. However, considering that site conditions have changed since the previous allocation, this reallocation should include all available equipment once again if it is to keep its optimal status. Thus, the whole resource allocation must be reorganized in order to optimally resume the execution of the project. This enhances the problem with a dynamic nonlinear feature, which must always be taken into account in earthworks design.

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, 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.

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 5.7.

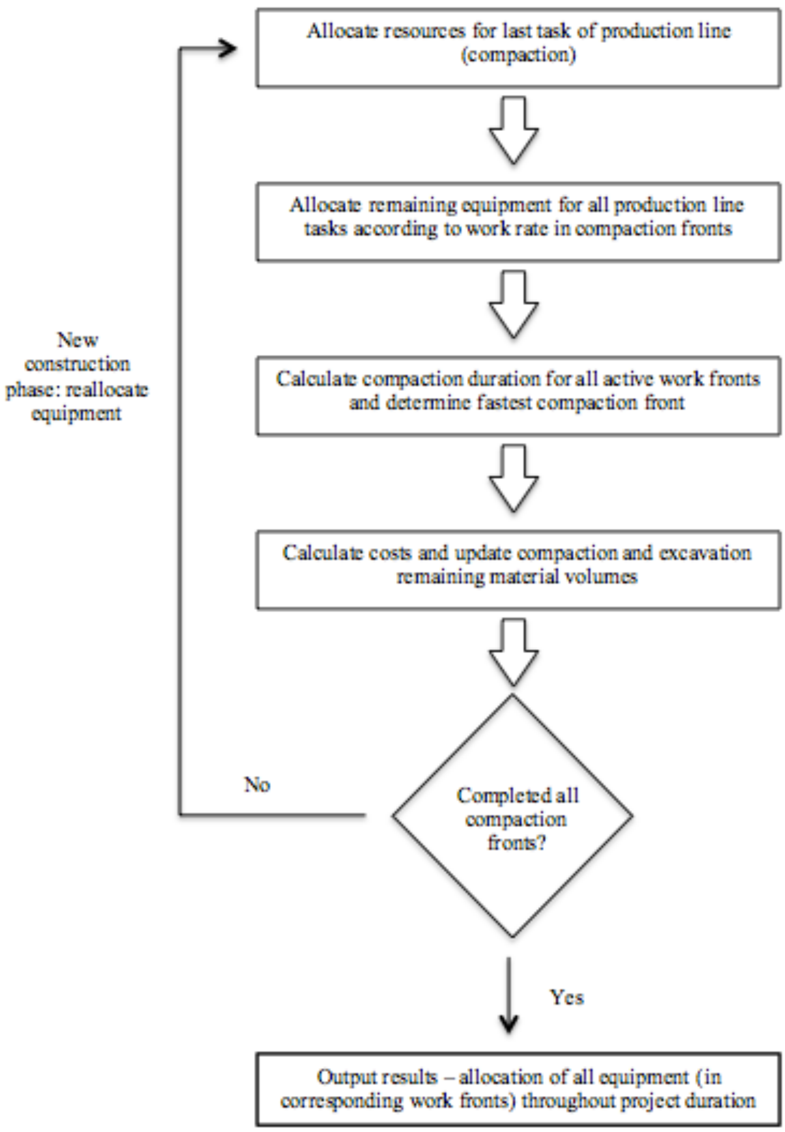


Figure 5.7 Algorithmic flow of dynamic earthwork resource allocation

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 5.1.

$$T_s = \frac{V_m}{Q_p} \tag{5.1}$$

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 5.2.

$$T_{pl} = \frac{V_c}{\min\left(Q_c, Q_s, Q_t, Q_e\right)} \tag{5.2}$$

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 equipment fuel or maintenance. In this case, the associated cost can be obtained as shown in Equations 5.3 (cost of a single piece of equipment, C_s) and 5.4 (costs per production line, C_{pl}).

$$C_s = \sum icost_i \times T_{pl} + \sum dcost_i \tag{5.3}$$

$$C_{pl} = \sum C_s \tag{5.4}$$

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 also discussed further ahead). 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.

Solution Quality Assessment

As referred in the previous subsection, 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 5.3.

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. 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 5.3. 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 m^3/h per compactor. This means that total productivity in front 1 would be 400 m^3/h, while front 2 would proceed at a rate of 800 m^3/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 m^3/h in front 1 and 800 m^3/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 plant is associated with a production line that can support its productivity. The methodology used for this allocation is described in the next subsection.

Table 5.3 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)

Having assembled the resulting production lines, the compaction time of each embankment front can be determined by applying Equation 5.2. 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 m^3/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 the next subsection. 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 5.3.

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 5.3).

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 5.3). 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 5.3, in which the required volume for embankment front 2 is now 4000 m^3, 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. The next subsection further details some aspects of this process, framing it in the context of a chromosomal representation adopted for an EMO algorithm.

Evolutionary Multi-Objective Optimization of Earthworks

The available resources (mechanical equipment) in an earthwork construction process can be allocated to each sequential task in the production lines that comprise it, 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.

The work presented in this paper is based on an evolutionary multi-objective optimization (EMO) algorithm, implemented in the form of a non-dominated sorting genetic algorithm II (NSGA-II). These types of evolutionary computation methods 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. In this context, 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 and Griffiths 2000; Xu et al 2011; Nassar and Hosny 2012). 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, 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: $g_1\ g_2\ g_3 \ldots g_C$, 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 5.8). This representation aims firstly to allocate each roller c_i to a compaction front $\in \{0, 1, 2, ..., F\}$ (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 5.8, 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	...	c_C	c_1	...	c_C	…	c_1	…	c_C
g_1	g_2	...	g_C	g_{C+1}	...	g_{2C}	…	$g_{C\times F-C+1}$	…	$g_{C\times F}$
Phase 1				Phase 2			…	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 5.8 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:

- **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.
- **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).
- **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 and 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 5.5 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.

Minimize $$\sum_{1}^{m_i} C_{s,i} \times x_i$$

Subject to $$\sum_{1}^{m_i} Q_{s,i} \times x_i \geq Q_{total,compaction}$$

$$x_i \geq m_i \tag{5.5}$$

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 productions lines, the compaction duration of all fronts is calculated according to Equation 5.2. 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 5.3 and 5.4.

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 5.6. 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.

$$F_{R,i} = F_{R,i} - \min\,(T_{pl,i}) \times \mathrm{Q}_i \tag{5.6}$$

Figure 5.9 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 5.8. Prior to the initial population generation step, it is possible to automatically run a simple linear programming model that carries out the material management, distributing the volumes of available geomaterial in the excavation fronts through the required volumes in each embankment front. This is carried out by default, although that information can alternatively be inputted into the system. The fitness function includes the repair methodology described above, 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 abovementioned formulae (i.e., Equations 5.2 and 5.4), 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 5.9 for each solution.

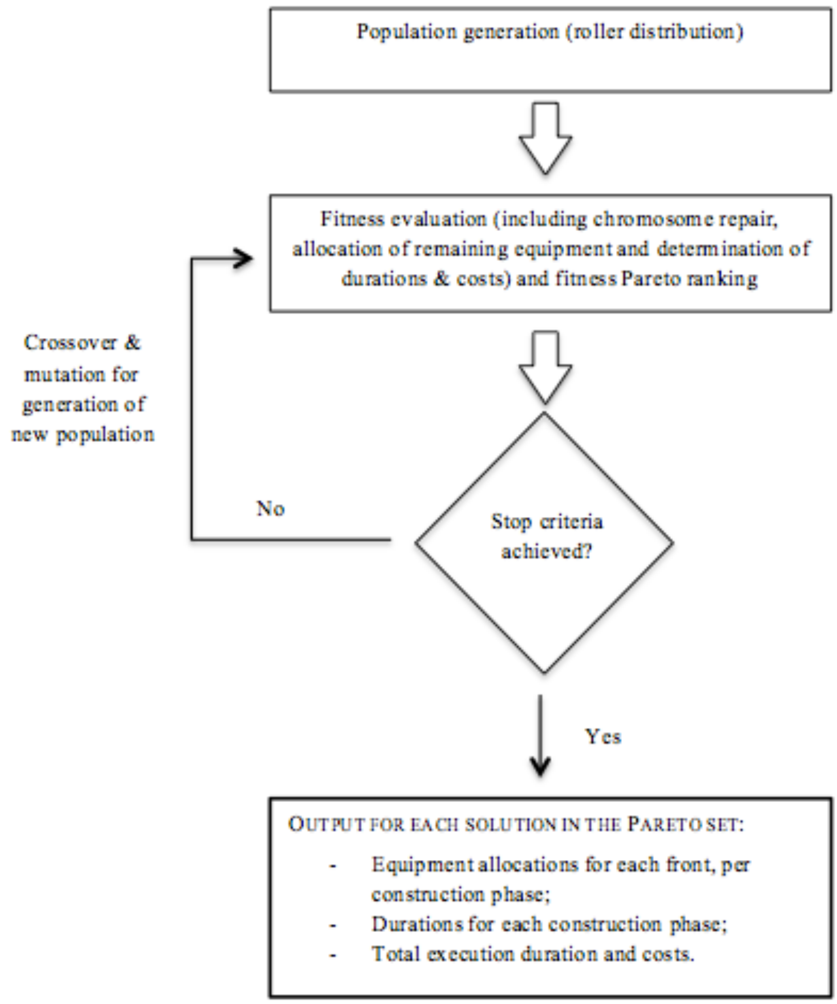


Figure 5.9 Algorithmic flow of the EMO algorithm.

5.3.4. Module Integration

The developed system is comprised of the equipment module (DM models) and spatial module (GIS spatial data), integrated into the multi-objective optimization module for equipment distribution in earthworks. The first two modules are directly or indirectly associated with estimating equipment productivity, while the third module carries out its optimal allocation. In this context, the DM models from the equipment module are used by the EMO in the optimization module to determine the productivity of compaction equipment for each embankment front. Concurrently, the GIS output from the spatial module is the basis for the estimation of transportation equipment productivity. The importance of these technologies to support the optimization module is paramount, as the productivity of these types of equipment (compaction and transportation) is heavily dependant on site conditions. This means that manufacturer specifications are often not enough to properly estimate compaction and transportation productivity in early phases of an engineering project (i.e., early design phase), when information is lacking.

In the original architecture proposed in Section 5.2, the user inputs the available equipment and spatial data into the equipment and the spatial modules, respectively, while the output is returned to the user from the optimization module (Figure 5.1). The equipment module inputs regard the material types and volumes in each work front, as well as the type and number of available equipment and associated specifications, while the spatial module inputs are related to creating the model of the construction area, as described in Section 5.3.2. These inputs are summarized in Table 5.4.

Table 5.4 Data from each module imported into the optimization module

Imported data (optimization module)	
Equipment module	*Spatial module*
- Material volumes required in embankment fronts and available in excavation fronts - Material type in each excavation front - Type and number of available equipment associated with each task (excavation, transportation, spreading and compaction) - Equipment direct and indirect costs and work rate (when not estimated by DM models)	- Optimal travel distance/time from each excavation front to each embankment front (OD cost matrix)

Initially, the spatial module was expected to automatically access data from the equipment module so as to obtain the equipment specifications that affect the workflow conditions, such as dimensions or maximum speed. However, since integration of the spatial module with the other modules was not accomplished due to the aforementioned mentioned incompatibilities, this information is presently inputted into the GIS software manually. In this way, these factors are taken in the final output of the spatial module. As such, after (manually) importing the GIS output into the optimization module, the latter uses this data in order to proceed with the allocation of mechanical resources, namely the transportation equipment, in function of the optimal haul distances or times, and depending on what was defined in the GIS software (user preference).

Regarding the equipment and optimization modules, initial integration attempts feature the loading of the DM models as an integral part of the optimization module fitness function, as shown in Figure 5.10. Nonetheless, this solution turned out to be demanding in terms of computation effort during the fitness evaluation of each solution. Since the fitness function is carried out several times, depending on the size of the population being maintained by the EMO, this methodology can amount to very high computational times.

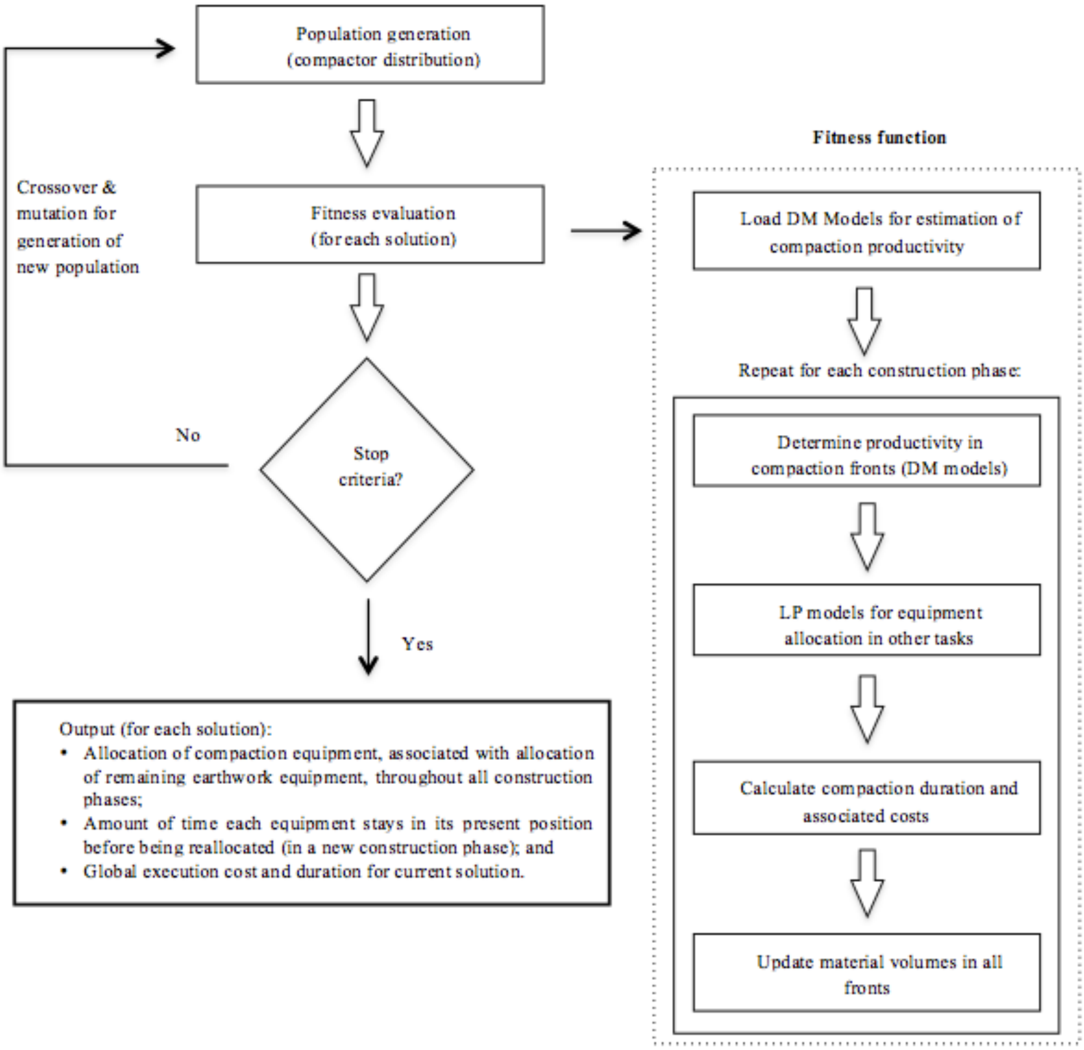


Figure 5.10 Initial integration of equipment and optimization modules (including manual inputs for GIS data)

A viable solution for this issue could be accomplished by loading the DM models prior to starting the optimization method (Figure 5.11). This is further facilitated through the development of a GUI for the optimization system. The development of a GUI can be achieved by converting the functions associated with the DM and optimization module into a *R* package, and using a GUI development package, such as *gWidgets* (Verzani 2014) or *RGtk2* (Lawrence and Lang 2014). This methodology allows the loading of the DM to coincide with the loading of the GUI. To be precise, the DM models are initiated while the package/system is loaded and prepared for use. In this way, the optimization system, namely the fitness function, simply has to take into account the already loaded DM models in order to determine the productivity of every compactor in each embankment front, depending on factors such as roller type, material classification and site conditions (e.g., layer thickness), as discussed in Chapter 2.

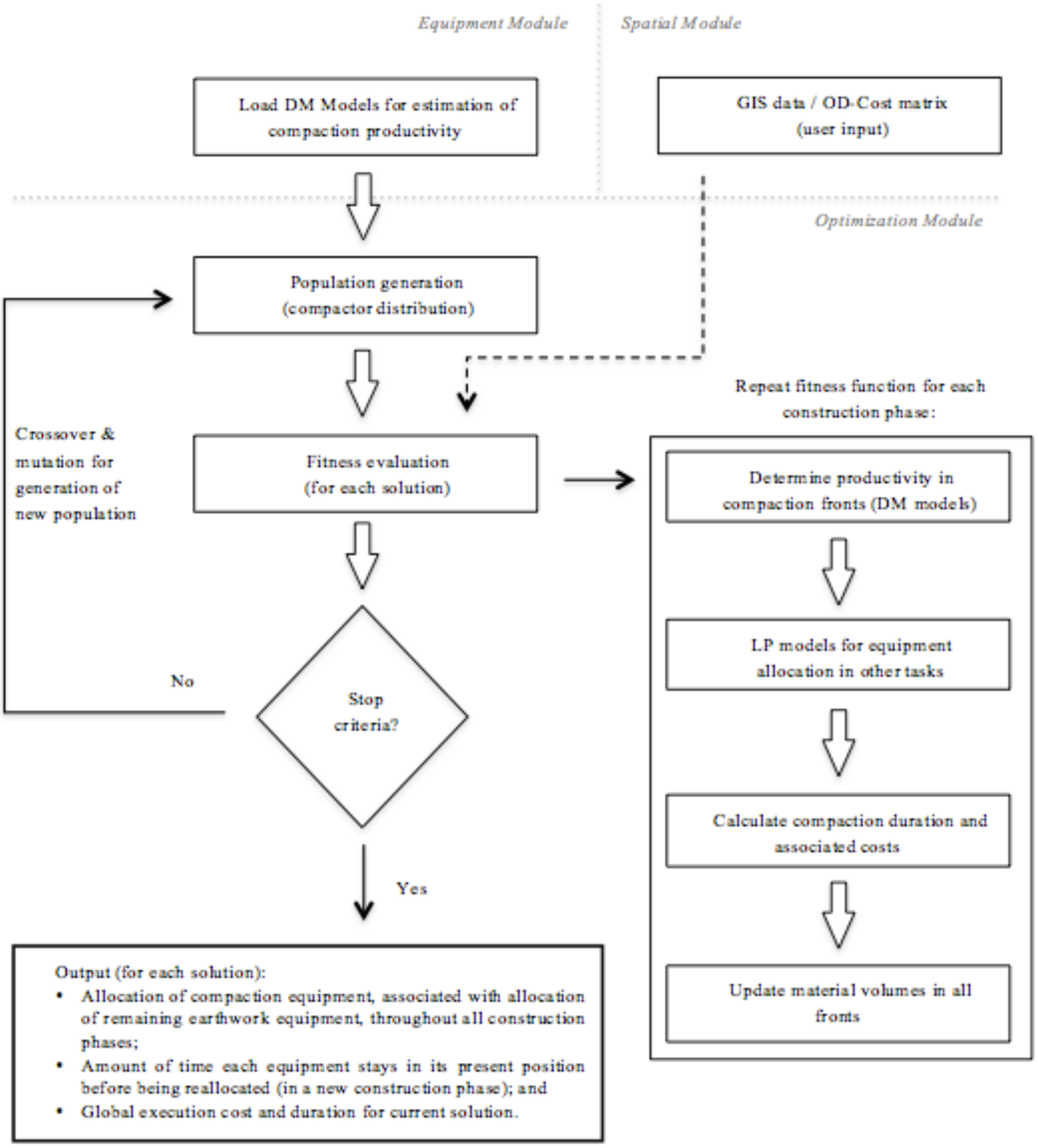


Figure 5.11 Algorithmic flow for the multi-criteria optimization system, including new integration methodology

Moreover, in order to further simplify and reduce the computation effort associated with the fitness evaluation of each individual by the optimization module, a compactor-material matrix can be generated during the insertion of inputs into the GUI. This matrix can be achieved if the user includes all the information regarding equipment and material specifications, as well as optional site conditions, such as layer thickness in each embankment front. Generating this matrix prior to initiating the optimization module (and after loading the package itself, including the DM models) facilitates the computation effort associated with the latter, as it will only have to access the matrix data in order to determine the productivity of a compaction equipment in its working conditions, as opposed to resorting to the DM model itself in every fitness run.

Note that the input of site conditions (e.g., embankment layer thickness) is referred to as optional during design phase of an engineering project, since this information may not yet be available. Whenever it happens, the system can simply assume the optimal values for these conditions, including layer thickness, since this information is available in the compaction guide used as a reference (GTR), as well as the DM models. As construction ensues and/or more information becomes available, the original user input can be updated, resulting in a new compactor-material matrix, which can be used in subsequent optimization runs.

In summary, integration is achieved in the optimization module itself, as the latter uses the output from the other two modules as input for the determination of productivity or allocation of resources. This is performed during the fitness function associated with each solution, which in turn is represented by a distribution of rollers throughout embankment fronts. The resulting allocation, which is carried out in the fitness function of every solution, allows for the determination of global construction cost and duration of each solution, thus permitting to assess the solution quality. Finally, the distribution of equipment in each solution is saved in memory, so that the user can consult the individual equipment allocation for the chosen Pareto solutions.

5.4. FINAL REMARKS

This Chapter proposes a novel EMO approach that addresses the whole earthwork construction phase and optimizes both cost and duration objectives. The approach includes a new representation of solutions where the EMO first allocates compaction equipment and then linear programming is used to distribute the remaining equipment (e.g., excavators and trucks). A Pareto approach based on a NSGA-II was selected as a basis for the allocation of the compaction equipment. The combination of capabilities of the technologies presented in Chapter 3, including artificial intelligence (in the form of evolutionary computation and data mining methods) and linear programming optimization, culminates in the proposal of the current system. In an attempt to adjust to the complex and dynamic reality associated with earthworks, as discussed in Chapter 2, the aim is to optimize the available resource allocation (represented by mechanical equipment) throughout the sequential tasks (namely excavation, transportation, spreading and compaction of geomaterials, as well as layer wettening, which is considered together with spreading) that comprise the earthworks process. In this framework, DM and GIS technologies support the optimization techniques by providing realistic estimates to the productivity of the available equipment given site conditions.

As will be seen in Chapter 6, corresponding to the application of the proposed system to a case study, the presented system shows the potential to surpass the limitations found in previous attempts at solving the earthwork optimization problem (Chapter 4). Besides the already mentioned innovative capabilities, one main advantage of this system is its flexibility, both in terms of adjusting to the ever-changing earthworks environment, and to the user's point of view regarding the several outputted Pareto optimal solutions that minimize final deadline and/or budget. In this sense, it can be regarded as a decision support system, which presents the user with several optimal trade-offs between the objectives, allowing the choice of the most appropriate for any given situation. Moreover, the system is expected to be able to include additional criteria to support the final decision, namely environmental aspects, which can be effortlessly assessed by the determination of carbon dioxide emissions in each solution (assuming the availability of this information, which is commonly supplied by equipment manufacturers).

Chapter 6

ROAD CONSTRUCTION: CASE STUDY

6.1. INTRODUCTION

The previous Chapters culminate on the proposal of the novel earthwork optimization system in Chapter 5. This innovative system aims to minimize global construction costs and durations by keeping a population of solutions comprised of a distribution of compaction equipment throughout the earthwork embankment fronts. To each of these solutions (compactor equipment distributions) corresponds a linear programming (LP) distribution of the remaining equipment, designed to homogeneize the productivity of each task preceding compaction. Besides minimizing construction costs and durations, by keeping a homogeneous distribution of productivity in each task, each piece of equipment is used to its maximum potential, eliminating any equipment idle time, and reducing environmental impact (e.g., reducing fuel usage and carbon emissions).

In this Chapter, the focus is on the application of the presented system. Considering that the system is innovative not only regarding the global improvement of the earthworks workflow, but also on account of the original solutions that have been adopted during its development, a twofold assessment is implemented:

- Assessment of the performance of the evolutionary multi-objective optimization (EMO) algorithm; and
- Analysis and discussion of results in the earthwork construction point of view.

In order to acquire a better insight on the obtained results, several non-dominated sorting genetic algorithm II (NSGA-II) runs were carried out on the same example. On the one hand, these runs were used to assess the performance of the EMO, in terms of required computational time and effort, as well as on the convergence of the algorithm. On the other hand, performing several runs allows for a robust analysis of the quality of results, since the adopted evolutionary computation techniques neither necessarily attain the same result on each run, nor guarantee the optimal solution.

6.2. DESCRIPTION OF CASE STUDY

In this application, the same real-world data related with the mechanical equipment of a Portuguese road construction site that was previously described in Chapter 5 (see Section 5.3.1) was adopted. The used database subset 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 m^3. The distances from excavation to embankment fronts vary between 100 m to nearly 4,000 m, stressing the need for proper resource management.

The available equipment of the construction company for these operations 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 and 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 approximately equals to 9.9h. Considering this setup, the final cost obtained for these work phases was close to 135,200 €.

The necessary inputs for the optimization system were retrieved from the available database and are summarized in Table 6.1. Once again, it is important to refer that the equipment and tasks associated with layer wettening are considered concurrently with the spreading task as far as productivity calculations are concerned. Equipment indirect costs and work rates derived from Caterpillar Performance Handbook (Caterpillar Inc., 1998) and Transportation Research Board NCHRP Report 744 (Skolnik et al 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 6.1 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	

6.3. COMPUTATIONAL EXPERIMENTS

6.3.1. EMO performance

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 5x5=25 genes each, defining the search space for this problem.

The default parameterization of the NSGA-II, as implemented in the *R* package *mco*, 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 conventional manual 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 256,642.5 seconds (around 70.73 hours).

To illustrate the multi-objective convergence, Figure 6.1 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 a non-convex near the bottom left region.

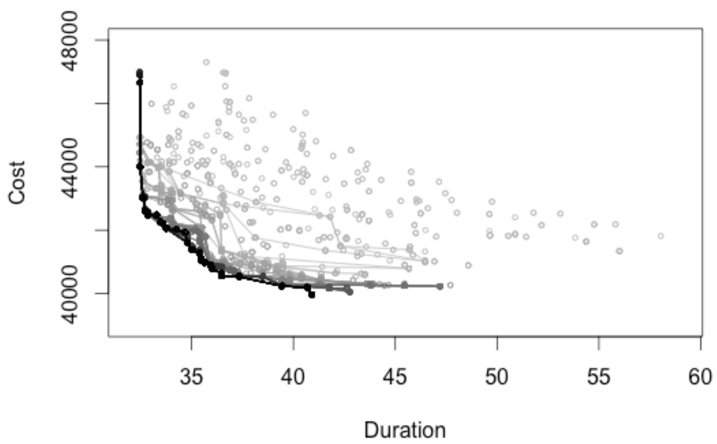


Figure 6.1 Example of the convergence of NSGA-II algorithm (*x*-axis in hours and *y*-axis in €)

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 et al 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 6.2, 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 receiver operating characteristic (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 is 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 6.2 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 close to 212,000 seconds (58.90 h).

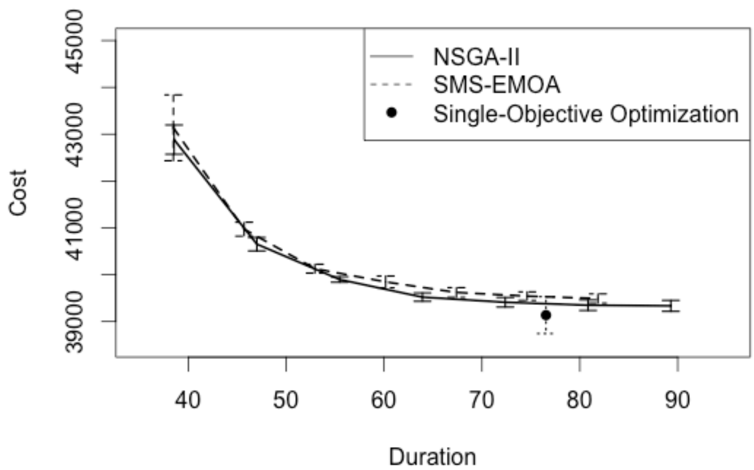


Figure 6.2 Comparison between results obtained by different optimization methods (x-axis is in hours and y-axis in €)

6.3.2. Earthwork workflow results

The proposed system results were compared against the original manual allocation performed by the construction company, as shown in Figure 6.3. 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 6.3, 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 135,200 € that was obtained in the original allocation. The original conventional manual allocation is far from the optimal solution, which can be explained by the lack of an efficient automated optimization and also by unforeseen delays.

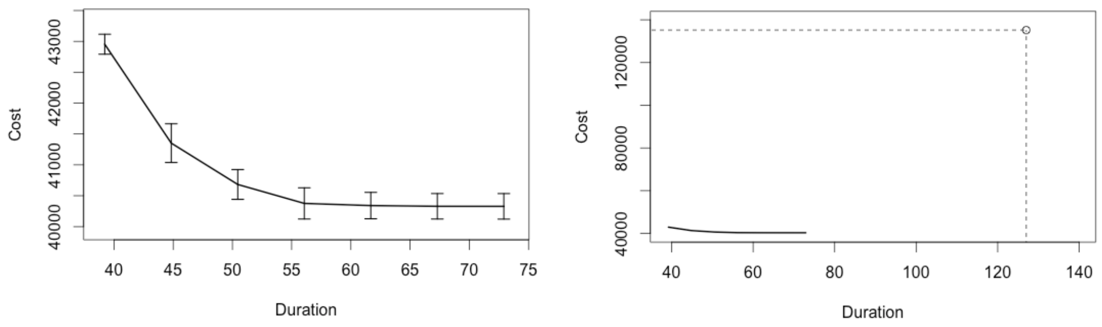


Figure 6.3 Optimization results (k=0.75) in terms of: left graph - vertically averaged Pareto front; and right - comparison between the optimized Pareto front (black curve) and the conventional manual allocation solution (circle point) (in both graphs x-axis is in hours and y-axis in €)

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 6.2. 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 m^3/h in production line 1 and 540 m^3/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.

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 indirectly result in minimization of unnecessary carbon emissions. This last aspect is very challenging to accomplish by conventional design methodologies.

Table 6.2 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	
Equipment distribution and work rate		*Equipment distribution and work rate*	
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 m3/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	
Equipment distribution and work rate		*Equipment distribution and work rate*	
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 m3/h (compaction)		Minimal production line work rate: 683 m3/h (compaction)	
Duration: 28.2 h		Duration: 15.6 h	
Cost: 12,718 €		Cost: 5,740 €	

Comparing the original setup with its optimized counterpart for production line 1, the conventional manual 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 plant. 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 plants. 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 manual 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 and 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 competiveness of the proposed system, additional NSGA-II experiments were conducted assuming a much lower efficiency factor of k=0.3 (Figure 6.4). 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 6.4). 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.

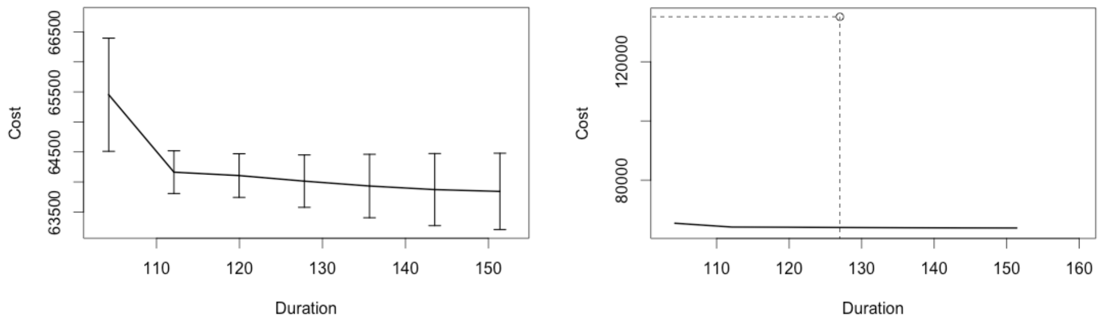


Figure 6.4 Optimization results (k=0.3) in terms of: left graph - vertically averaged Pareto front; and right - comparison between the optimized Pareto front (black curve) and the conventional manual allocation solution (circle point) (in both graphs x-axis is in hours and y-axis in €).

6.4. FINAL REMARKS

Bearing in mind how earthworks comprise a high percentage of costs and durations in many Civil Engineering projects, namely road, railway and airport construction, their optimization should always be a priority. Given this, any reduction on that percentage will have a significant effect on the final cost and duration of the whole construction project. In this context, 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.

The novel system proposed in Chapter 5 was applied in a case study using real-world data from a Portuguese construction site, showing that the EMO system is quite competitive when compared with conventional design, and that it can be easily adapted to dynamic changes that are inherent to earthworks constructions. In fact, for this case study, a high impact on would be achieved by the implementation of this system, as results indicate a reduction of approximately 50% in construction cost and duration when compared with the originally adopted solution (achieved via conventional manual design). In an attempt to gain some insight on the actual efficiency of the equipment plant during the original construction, additional computational experiments were carried out under extremely lower efficiency ratios (k=0.3, corresponding to 18 minutes of actual work per hour). However, even in this case, the optimized solutions still slightly outperform the manual allocation solution cost-wise. Naturally, these results do not take into account the possible delays and costs associated with unpredictable events and obstacles that occur during construction (e.g., equipment malfunction). However, the system features the flexibility to deal with these issues, since it allows for the user to easily rerun the optimization procedure with an updated set of conditions and constraints (e.g., less available equipment), which outputs a new set of optimal allocation solutions.

The computational experiments presented in this Chapter not only bring forth the potential of the system, but also identify some significant of the limitations of conventional manual earthworks design, in particular where the production line equipment is either significantly above the required work rate requirements (incurring in unnecessary costs) or below it (resulting in idle times and low efficiency ratios). Moreover, it was possible to verify the capability of the proposed system to distribute equipment in a relatively homogeneous way (when compared to conventional design), while minimizing costs and durations, which was one of the goals of this research.

Chapter 7

CONCLUSIONS AND FUTURE WORK

7.1. SYNOPSIS

This thesis describes a complete research and development cycle, starting from the definition of a relevant problem (i.e., how to achieve an optimal allocation of equipment throughout an earthworks project), to the implementation and application of an innovative solution, capable of dealing with the high complexity of the referred problem. As such, this thesis is divided into six chapters. The first five include the description of both the earthworks optimization problem and the developed system, culminating with the depiction of a tool capable of supporting the design and allocation of equipment in earthwork constructions. The final part comprises the results of the application of this system in a case study associated with a Portuguese earthwork construction site.

As such, Chapter 2 accounts for the earthwork construction background, focussing on its definition as an optimization problem. The significant consequences on the final costs and duration of the optimization of projects that include earthworks, such as road, railway or airway construction, are emphasized. Moreover, particular stress is laid on how conventional manual design is usually carried out for this type of construction, including the present concerns related with earthworks design. Understanding the relevance of these aspects is essential to the development of a solution capable of dealing with the earthwork optimization problem.

The first step to achieve this goal must be the study of the relevant technologies that show potential to deal with complex and dynamic environments, such as the optimization of earthwork tasks. This is the main focus of Chapter 3, in which the artificial intelligence area, namely soft computing techniques (i.e., data mining and metaheuristics), as well as geographic information systems, are established as the tools with full potential to deal with the problem under study.

This notion is confirmed in Chapter 4, in which a study regarding previous attempts on using the techniques brought up in Chapter 3 is carried out in order to address the issues associated with the problem at hand (described in Chapter 2). The study is not limited to a literature research on the subject, as most of these attempts are concretized in the form of systems that aim to either support or optimize the design of earthworks. Subsequently, this allows for a more complex analysis of each system, sorting them into general types with similar characteristics and analysing their strengths and limitations, as a function of the technology they are based upon, and the part of the problem they address. However, one of the conclusions drawn from this study is that most of these systems are limited in more than one way. In fact, many do not explore the full advantages of performing multi-objective optimization.

These limitations constitute a gap, for which further development and innovation is required. Given this prospect, allied with the fact that the necessary technology for dealing with the complexity of this problem is available nowadays, a novel earthwork optimization system is proposed and described throughout Chapter 5. The technologies mentioned in Chapter 3 are integrated into one system in order to surpass the limtations of the systems studied in Chapter 4, while simultaneously gathering all their strengths. The system is divided into three modules, each one associated with different technologies, which communicate and interact together to generate practical and flexible solutions, with the potential of supporting both design and construction phases of an earthwork project.

Finally, the potential of this system is demonstrated in Chapter 6, in parallel with its application in a case study using real-world data. Since this is a multi-disciplinary work, including areas such as civil engineering, information systems, and geographic information systems, this application is performed in two different approaches. On the one hand, the application is aimed at assessing the performance of the used optimization algorithms, as well as the viability of the innovative solutions and methodologies to deal with the complexity of the problem in the information systems area. On the other hand, the analysis of results in terms of improvement of the earthworks process in the civil engineering point of view is also a fulcral point, showing the competitiveness and flexibility of the system when compared with conventional design.

7.2. DISCUSSION

Innovation can generally be defined as changing or creating more effective processes, products and ideas. It is related to concepts like unexpected connections, new ideas, or even original notions stemming from imagination. Businesses that innovate create more efficient work processes and have better productivity and performance. Indeed, innovation can be seen as a competitive advantage to grow and adapt a business to an area or marketplace.

Being innovative does not necessarily mean inventing. Innovation can mean changing a business model, or adjusting to environmental changes, in order to deliver better products or services. Successful innovation, based on the creation of a culture of innovation, should be an in-built part of a business strategy, leading the way in creative thinking and problem-solving. In this context, technological innovation is often defined as combining technologies from different environments in order to create an original method or technology that performs better than previous ones in a new environment.

Under this premise, this work is related to the integration of several different technologies from different fields into a novel system capable of solving a dynamic and complex problem in the earthworks area. The combination of technologies, such as data mining, modern optimization and geographic information systems, has proven to originate a powerful tool to support the design and management of any construction involving earthwork tasks, as demonstrated in Chapters 5 and 6. Innovation in the developed system derives not only from the use of diverse technologies from different areas, but also from some of the solutions and methodologies followed to achieve the proposed goals. In actual fact, the proposed evolutionary multi-objective approach, which addresses the whole earthwork construction phase and optimizes both cost and duration objectives, includes a new representation of solutions where the compaction equipment is allocated first, and then linear programming is used to distribute the remaining equipment (e.g., excavators and trucks). This innovative representation enables the system to globally optimize the whole earthworks process, as opposed to optimizing only partial tasks or construction aspects, which is verified in most previous attempts at solving this optimization problem.

Indeed, considering the analysis of previous earthwork optimization systems carried out in Chapter 4, most approaches simplify the complex reality of earthwork construction by separating the process into several independent parts, which are optimized individually. While this is understandable as an attempt to deal with the complex and dynamic nature of the problem, it neglects the advantages of global optimization. For example, when addressing each earthwork task individually, it may look advantageous to use the minimum amount of equipment in each step so as to minimize costs. However, when looking at the whole earthwork process as a sequential and interdependent process, it can, in many cases, be advantageous to add more pieces of equipment to a critical task, since this will have an important impact on final project duration and/or costs. Another common limitation among previously developed systems is related to either focusing on single-objective optimization or attempting to carry out multi-objective optimization by transforming multiple objectives into a single objective, as discussed in Chapters 4 and 5. The latter can be performed by preference-based methods, such as weight-based optimization. Although these can reduce computational effort requirements, they output only a single optimal solution, forcing the user to assign an importance (or weight) to each objective. These methodologies ultimately overlook the flexibility and potential of multi-objective optimization methods, such as a Pareto approach. In a Pareto approach, which has been adopted in this work, the output is not a single solution, but a set of optimal trade-offs between conflicting objectives (project cost and duration). This allows the user to choose the solution that best fits present construction conditions and criteria, enhancing the system with the flexibility to adapt to the dynamic and uncertain environment inherent to earthworks. Finally, during the application of the system in Chapter 6, the system has been shown to surpass some limitations of the conventional manual earthworks design, demonstrating its competitiveness and adaptation capabilities to dynamic changes in earthwork construction.

Consequently, the final result is a system that can be used in every part of an earthworks project, from planning and design phases to construction phase. During planning and design phases, the lack of reliable data on site conditions, essential to support a proper earthworks design, can be compensated by the developed data mining models, which are integrated with the optimization system. Given their learning potential, these models facilitate the estimation of unknown or hard to predict parameters, such as equipment productivity, allowing for several optimization and allocation analysis, in order to support decision-making and proper design. Furthermore, the fact that a Pareto approach has been adopted to deal with the multi-objective feature of the problem results in several solutions being output by the system for equipment allocation. Hence, the designer has several options to choose from, enhancing the planning of earthworks with the flexibility to adjust to different scenarios. This feature is also noteworthy as far as construction phases are concerned, where dynamic and unpredictable situations can occur at any given time. Common situations that occur in earthworks, such as alteration of site conditions (e.g., concluding that the geomaterial quality is not as high as originally expected, or transportation routes becoming unusable during construction) or equipment restrictions (e.g., malfunctioning of some equipment, making it unavailable for further use either temporarily or permanently), can be easily dealt with by adjusting the previously developed models to include the new conditions or restrictions. This will result in rerunning the optimization process, which, in turn, will output a new series of Pareto optimal solutions. Given that, at that time, the available budget can be substantially limited, or the final deadline considerably closer, the possibility to choose from several options may be an important advantage, since the designer can pick the one that best fits current restrictions. This process can be easily reiterated at any time during construction phase, in order to readjust to new conditions and adapt to unforeseen situations.

Finally, it is important to highlight that, although project cost and duration, as well as carbon emissions (indirectly), were the main minimization objectives mentioned and approached throughout this work, it is viable to include other objectives and points of view in the optimization procedure. As a matter of fact, environmental impact in this work is mainly addressed in the form of carbon emissions, which are naturally minimized once equipment usage is optimized to its full potential. However, any factor that can be mathematically quantified and related to the earthworks allocation problem can be used as an optimization objective. Environmental impact does not have to be limited to carbon emission. In fact, as discussed in Chapter 2 (Section 2.4), environmental concerns in earthwork construction are much wider, ranging from waste minimization to water conservation during construction. Another interesting point of view that could be considered would be the social and economic aspect of an earthwork construction in a given region, which can even be correlated to the associated environmental impacts, linked with sustainability principles. Indeed, considering how the use of a higher amount of available machinery in a construction project can lead to an increase in job opportunities for that region, as opposed to minimizing the number of active equipment, so as to diminish the environmental impact of carbon emissions. Naturally, increasing the number of job opportunities will also have a significant impact on the final cost of the whole project. However, it can also be advantageous for the economic aspects of the region in which the construction project is executed. It is obvious that increasing the number of conflicting optimization objectives will result in an increase in the optimization complexity, meaning that further study is required to understand how high the this increase in complexity really is when adding a new optimization objective. As such, this aspect is also mentioned in the next section, corresponding to future works (see Section 7.3).

7.3. FUTURE WORKS

Having addressed a complex problem, to which an innovative solution was developed, this work has naturally opened up various research possibilities where further innovation can ensue:

- As referred throughout the description of the development of the system, in Chapter 5, the full integration of all technologies and tools would enhance not only the system's capabilities, but also its user-friendliness. Since the integration of geographic information system technologies was not fully achieved throughout the development of this project, future growth should be based either on subsequent versions of the presently used tools, or on migrating/expanding the system towards the use of new tools altogether.
- Given the timespan of the project, some simplifications of reality were considered during the development of the prototype system. As a consequence, its future development, allowing for a superior adjustment and modeling of reality, should be thought through. A clear example might be related to better grasping some aspects regarding site conditions, particularly space restrictions in work fronts, which play an important role in limiting the maximum amount of active equipment in a single work front. Indeed, an interesting application will be in urban excavation for buildings,

parking lots and underground, among others. Another important aspect to consider concerns situations in which an embankment front is being fed different types of geomaterial from different excavation fronts. In these cases, constructive aspects such as using a particular material in the lower layers, and other in the upper layers, could be taken into consideration by making a sensible decision on the timing or the order of the excavation process in each front. This development would increase the flexibility of the system to deal with this practical issue, further supporting the decision of designers and engineers.

- The development of the system focuses mainly on the issues of allocation of available resources (the number of machines applied), the selection of the best equipment fleet (the selection of machines to be applied) and proper application of machines in terms of haul/return routes. While these are important factors to achieve maximum productivity and minimum cost, other aspects like loading and dumping configurations, or space optimization in work fronts (e.g., where to place machinery in a work front, so as to facilitate and maximize the interaction between equipment) should as well be taken into account for a complete and integrated optimization.
- Gathering more knowledge on the optimization parameters that can output the best genetic algorithm results, as well as on the exploration of other multi-objective optimization methods, such as Strength Pareto Evolutionary Algorithm 2 (SPEA-2) or S-Metric Selection Evolutionary Multi-objective Optimization Algorithm (SMS-EMOA), are also important aspects to be addressed in future work. Moreover, as referred in the end of the previous section, more knowledge regarding the addition of further optimization objectives could help understand how far the optimization of earthworks can go.
- Bearing in mind the increasing importance of sustainability in construction in recent years, it is imperative to explore the potential development of a new module (or the enhancement of the already developed modules) towards taking this aspect into account. The module should be able to objectively determine the sustainability index associated with an earthwork project, accounting, for instance, for carbon emissions and material treatment procedures.
- Finally, the application/validation of the system for a practical project carried out in real time can provide the essential insight regarding the development of new functionalities and the improvement of current ones.

BIBLIOGRAPHY

AbouRizk SM, Hajjar D (1998) A framework for applying simulation in construction. Can J Civ Eng 25:604–617.

Alkass S, Harris F (1988) Expert system for earthmoving equipment selection in road construction. J Constr Eng Manag 114:426–440.

Alkass S, Harris F (1991) Development of an integrated system for planning earthwork operations in road construction. Constr Manag Econ 9:263–289.

Askew WH, Al-jibouri SH, Mawdesley MJ, Patterson DE (2002) Planning linear construction projects: automated method for the generation of earthwork activities. Autom Constr 11:643–653.

British Standards Institution (1981) Code of practice for Earthworks.

Brown ML, Kros JF (2003) Data mining and the impact of missing data. Ind Manag Data Syst 103:611–621.

Caterpillar Inc. (2011) Caterpillar Performance Handbook Edition 29.

Chang F, Jiang B, Tang N (2000) The development of intelligent decision support tools to aid the design of flexible manufacturing systems. Int J Prod Econ 65:73–84.

Chapman P, Clinton J, Kerber R, et al (2000) CRISP-DM 1.0 Step-by-step data mining guide.

Cheng F, Wang Y, Ling X (2010) Multi-Objective Dynamic Simulation-Optimization for Equipment Allocation of Earthmoving Operations. Constr Res Congr 328–338.

Cheng F, Wang Y, Ling XZ, Bai Y (2011) A Petri net simulation model for virtual construction of earthmoving operations. Autom Constr 20:181–188. doi: 10.1016/j.autcon.2010.09.015

Cheng T, Feng C, Chen Y (2005) A hybrid mechanism for optimizing construction simulation models. Autom Constr 14:85–98.

Cortez P (2010a) Package “rminer.”

Cortez P (2010b) Data Mining with Neural Networks and Support Vector Machines using the R/miner Tool. Adv Data Min - Appl Theor Asp 6171:572–583.

Cortez P (2014) Modern optimization with R, Use R! Springer

Cortez P, Marques R, Gomes Correia A (2008) Artificial intelligence applied to compaction rules and management. Int. Semin. Interact. Soil-Rail Track High Speed Railw.

Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 6:182–197.

Denègre J, Salgé F (1996) Les systèmes d'information géographique. Presses Universitaires de France, Paris, France

Dorigo M, Gambardella LM (1997) Ant colony system: A cooperative learning approach to the traveling salesman problem. Evol Comput IEEE 1:1–24.

Drudzel MJ, Flynn RR (2002) Decision Support Systems. Encycl. Libr. Inf. Sci.

Dutta A (1996) Integrating AI and optimization for decision support : A survey. Decis Support Syst 18:217–226.

Edwards DJ, Griffiths IJ (2000) Artificial intelligence approach to calculation of hydraulic excavator cycle time and output. Min Technol 109:23–29.

Emmerich M, Beume N, Naujoks B (2005) An EMO algorithm using the hypervolume measure as selection criterion. Evol Multi-Criterion Optim 62–76. doi: 10.1007/978-3-540-31880-4_5

ESRI (2011) ArcGIS Desktop: Release 10.

European Construction Technology Platform (2005) Strategic Research Agenda for the European Underground Construction.

Farinha F, Bento J, Blockley D (1995) A Systems Approach to the Computer Aided Design of Reinforced Concrete Structures. Educ Pract Promot Comput Methods Eng 1:381–387.

Fawcett T (2006) An introduction to ROC analysis. Pattern Recognit Lett 27:861–874. doi: 10.1016/j.patrec.2005.10.010

Fayyad U, Piatetsky-Shapiro G, Smyth P (1996a) The KDD process for extracting useful knowledge from volumes of data. Commun ACM 39:27–34. doi: 10.1145/240455.240464

Fayyad U, Piatetsky-Shapiro G, Smyth P (1996b) From Data Mining to Knowledge Discovery in Databases. Am Assoc Artif Intell 17:1–18.

Glover F (1986) Future paths for integer programming and links to artificial intelligence. Comput Oper Res 13:533–549.

Glover F, Laguna M (1997) Tabu Search. Kluwer Academic Publishers

Göktepe AB, Lav AH, Altun S, Altintas G (2008) Fuzzy decision support system to determine swell/shrink factor affecting earthwork optimization of highways. Math Comput Appl 13:61–70.

Gomes Correia A, Brandl H, Magnan J-P (2014) Earth and rockfill embankments for roads and railways : what was learned and where to go. XV Danube-European Conf. Geotech. Eng. (Keynote Lect.)

Gomes Correia A, Cortez P, Tinoco J, Marques R (2012) Artificial Intelligence Applications in Transportation Geotechnics. Geotech Geol Eng 31:861–879.

Gomes Correia A, Magnan J-P (2012) Trends and challenges in earthworks for transportation infrastructures. Adv. Transp. Geotech. 2. Taylor & Francis Group, pp 1–12

Google Inc. (2009) Google Earth (Version 5.1.3533.1731).

Greater London Authority, London Councils (2006) The Control of Dust and Emissions from Construction and Demolition Best Practice Guidance. Publ. by Gt. London Auth.

Groth R (2000) Data Mining: Building Competitive Advantage. Prentice Hall, USA

Gutiérrez Puebla J, Gould M (1994) SIG: sistema de información geográfica. Editorial Síntesis

Haeser G, Gomes-Ruggiero M (2008) Aspectos Teóricos de Simulated Annealing e um Algoritmo Duas Fases em Otimização Global. Tendências em Matemática Apl e Comput 9:395–404.

Hastie T, Tibshirani R, Friedman J (2009) The Elements of Statistical Learning: Data Mining, Inference, and Prediction., Second Edi. Springer, New York

Hebb DO (1949) The Organization of Behavior. John Wiley & Sons, New York

Hevner AR, March ST, Park J, Ram S (2004) Design Science in Information Systems Research. MIS Q 28:75–105.

Heywood I, Cornelius S, Carver S (2011) An Introduction to Geographical Information Systems. Pearson

Hola B, Schabowicz K (2010) Estimation of earthworks execution time cost by means of artificial neural networks. Autom Constr 19:570–579. doi: 10.1016/j.autcon.2010.02.004

Holland JH (1975) Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor, Michigan

Holsapple C (1977) Framework for a Generalized Intelligent Decision Support System. Purdue University

Jha MK (2012) Metaheuristic Applications in Highway and Rail Infrastructure Planning and Design: Implications to Energy and Environmental Sustainability, 1st Editio.

Kataria S, Samdani SA, Singh AK (2005) Ant Colony Optimization in Earthwork Allocation. Int Conf Intell Syst 1–9.

Kennedy J, Eberhart RC (1995) Particle Swarm Optimization. Proc. IEEE Int. Conf. Neural Networks. pp 1942–1948

Kim E, Jha MK, Son B (2005) Improving the computational efficiency of highway alignment optimization models through a stepwise genetic algorithms approach. Transp Res Part B 39:339–360. doi: 10.1016/j.trb.2004.06.001

Kim S-K, Russell JS (2003) Framework for an intelligent earthwork system Part I . System architecture. Autom Constr 12:1–13.

Kim Y, Street WN (2004) An intelligent system for customer targeting: a data mining approach. Decis Support Syst 37:1619–1636.

Kirkpatrick S, Gelatt CD, Vecchi MP (1983) Optimization by simulated annealing. Science (80) 220:671–680.

Kotdwala SJ, Basheer IA (1994) Quality management of earthwork using knowledge-based systems. Comput Civ Eng 426–429.

Lawrence M, Lang DT (2014) Package “RGtk2”: R bindings for Gtk 2.8.0 and above.

Li L, Tang H, Wu Z, et al (2004) Data mining techniques for cancer detection using serum proteomic profiling. Artif Intell Med 32:71–83.

Liao S, Chu P-H, Hsiao P-Y (2012) Data mining techniques and applications – A decade review from 2000 to 2011. Expert Syst Appl 39:11303–11311. 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. J Constr Eng Manag. doi: http://dx.doi.org/10.1061/(ASCE)CO.1943-7862.0000940

London Health Comission, London Development Agency (2004) Rapid Health Impact Assessment of the Proposed London Olympic Games and Their Legacy. Final Report. London, UK

Luo W, Liu Q, Hu Z, Qiu Y (2008) The Simulation Study on Dynamic Optimization of Hydropower Project Earthwork Allocation System Based on Petri Net. 4th Int Conf Wirel Commun Netw Mob Comput 1 – 4. doi: 10.1109/WiCom.2008.1836

Mandow L, Pérez-de-la-Cruz J-L (2004) Sindi: an intelligent assistant for highway design. Expert Syst Appl 27:635–644. doi: 10.1016/j.eswa.2004.06.005

Marques R (2007) Desenvolvimento de um sistema pericial para apoio à optimização da compactação de geomateriais em vias de comunicação. University of Minho

Marques R, Gomes Correia A, Cortez P (2008) Data Mining Applied to Compaction of Geomaterials. Eight Int. Conf. Bear. Capacit. Roads, Railw. Airfields

Martin D (1991) Geographic information systems and their socioeconomic applications. Routledge, London, UK

Marzouk M, Moselhi O (2002a) Optimizing earthmoving operations using computer simulation.

Marzouk M, Moselhi O (2002b) Selecting Earthmoving Equipment Fleets Using Genetic Algorithms. In: Yucesan E, Chen C-H, Snowdon JL, Charnes JM (eds) Proc. 2002 Winter Simul. Conf. Montreal, Canada, pp 1789–1796

Matsatsinis N, Siskos Y (2002) Intelligent support systems for marketing decisions.

Mawdesley MJ, Askew WH, Al-jibouri SH (2004) Site layout for earthworks in road projects. Eng Constr Archit Manag 11:83–89.

Mawdesley MJ, Cullingford G, Haddadi A (1988) Site Layout and Resource Scheduling - An Approach to Modelling Movement Around Sites. 5th Int. Symp. Robot. Constr. Tokyo, Japan, pp 391 – 399

McCulloch WS, Pitts W (1943) A Logical Calculus of the Ideas Immanent in Nervous Activity. Bull Math Biophys 5:115–133.

Mersmann O (2012) Package “emoa”: Evolutionary Multiobjective Optimization Algorithms.

Mersmann O, Trautmann H, Steuer D, et al (2014) Package “mco”: Multiple Criteria Optimization Algorithms and Related Functions.

Miao K, Li L, Yang X-L, Huo Y-Y (2009) Ant colony optimization algorithm for vertical alignment of highways. ASCE Geotech Spec Publ 99–108.

Miao K, Sun X, Li L (2011) A roadbed earthwork allocation model based on ACO algorithm. Appl Mech Mater 44-47:3483–3486.

Michalewicz Z (1999) Genetic Algorithms + Data Structures = Evolution Programs, 3rd Editio. Springer

Michalewicz Z, Fogel DB (2000) How to Solve It: Modern Heuristics, 2nd Editio. Springer

Michalewicz Z, Schmidt M, Michalewicz M, Chiriac C (2007) Adaptive Business Intelligence. Springer-Verlag

Morgan W, Peterson L (1968) Determining shovel-truck productivity. Min Eng December:76–80.

Moselhi O, Alshibani A (2007) Crew optimization in planning and control of earthmoving operations using spatial technologies. J Inf Technol Constr 12:1–17.

Moselhi O, Alshibani A (2009) Optimization of earthmoving operations in heavy civil engineering projects. J Constr Eng Manag 135:948–954.

Muenchen RA, Hilbe JM (2010) R for Stata users. Springer-Verlag

Musaev A (2004) Analytic information technologies in oil refinery. Expert Syst Appl 26:81–85.

Nahar S, Sahni S, Shragowitz E (1986) Simulated annealing and combinatorial optimization. Proc 23rd Conf Des Autom 293–299.

Naoum S, Haidar A (2001) A hybrid knowledge base system and genetic algorithms for equipment selection. Eng Constr Archit Manag 7:3–14.

Nassar K, Hosny O (2012) Solving the Least-Cost Route Cut and Fill Sequencing Problem Using Particle Swarm. J Constr Eng Manag 138:931–942.

Naval P (2013) Package “ mopsocd .”

Olympic Delivery Authority (2007) Sustainable Development Strategy. London, UK

Parente M, Gomes Correia A, Cortez P (2014) Artificial Neural Networks Applied to an Earthwork Construction Database. In: Toll D, Zhu H, Osman A, et al (eds) Second Int. Conf. Inf. Technol. Geo-Engineering. IOS Press, Durham, UK, pp 200–205

Pearson K (1908) On the generalized probable error in multiple normal correlation. Biometrika 6:59–68.

Petri CA (1966) Fundamentals on the description of discrete processes. 3rd Colloq. Autom. Theory

Peuquet DJ, Marble DF (1990) Introductory Readings in Geographical Information Systems. Taylor & Francis, London, UK

Portela E, Bento J (2001) Dam Safety: the search for innovative tools. Int Water Power Dam Constr 53:26–31.

Power DJ (2002) Decision Support Systems: Concepts and Resources for Managers. Quorum Books

Quantum GIS Development Team (2015) Quantum GIS Geographic Information System.

R Development Core Team (2011) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Rinehart R V., Mooney MA (2008) Instrumentation of a roller compactor to monitor vibration behavior during earthwork compaction. Autom Constr 17:144–150.

Rosenblatt F, Cornell Aeronautical Laboratory (1958) The perceptron: A theory of statistical separability in cognitive systems. Cornell Aeronautical Laboratory, Buffalo, N.Y.

Rowley J, Gilbert P (1989) The Market for Land Information Services, Systems and Support. In: Shand, Moore (eds) Assoc. Geogr. Inf. Yearb. - 1989. Taylor & Francis, London, UK, pp 95–91

Santos MF, Azevedo C (2006) Data Mining - Descoberta de Conhecimento em Bases de Dados. FCA Editora de Informática, Lisbon, Portugal

Schabowicz K, Hoła B (2008) Application of artificial neural networks in predicting earthmoving machinery effectiveness ratios. Arch Civ Mech Eng 8:73–84. doi: 10.1016/S1644-9665(12)60123-X

SETRA, LCPC (2000) Guide des Terrassements Routiers - Réalisation des Semblais et des Couches de Forme.

Shi JJ (1999) A neural network based system for predicting earthmoving production. Constr Manag Econ 17:463–471.

Silva J, Azenha M, Gomes Correia A, Ferreira C (2013) Continuous stiffness assessment of cement-stabilised soils from early age. Géotechnique 63:1419–1432.

Skolnik J, Brooks M, Oman J (2013) NCHRP Report 744 - Fuel Usage Factors in Highway and Bridge Construction. Washington, D.C.

Sol HG (1985) Expert systems and artificial intelligence in decision support systems. Proc. Second Mini Euroconference

Sujitjorn S, Kulworawanichpong T, Puangdownreong D, Areerak K (2006) Adaptive Tabu Search and Applications in Engineering Design. Integr. Intell. Syst. Eng. Des. pp 233–257

Tam CM, Tong T, Tse S (2002) Artificial neural networks model for predicting excavator productivity. J Eng Constr Archit Manag 9:446–452.

The MathWorks Inc. (2014) MATLAB Release 2014a.

Tinoco J, Gomes Correia A (2013) Prediction of Soil-Cement Mechanical Properties Using EC2 Approach. Adv Mater Res 779-780:286–289. doi: 10.4028/www.scientific.net/AMR.779-780.286

Tinoco J, Gomes Correia A, Cortez P (2011) Application of Data Mining Techniques in the Estimation of Mechanical Properties of Jet Grouting Laboratory Formulations over Time. Soft Comput. Ind. Appl.

Transportation Research Board, National Research Council (1990) Guide to Earthwork Construction.

Turban E, Aronson J, Liang T (2004) Decision Support Systems and Intelligent Systems. Prentice Hall, New Jersey

Turban E, Sharda R, Delen D (2007) Decision support and business intelligence systems, 9th Editio. Pearson

Vapnik (1998) Statistical Learning Theory, volume 1851. John Wiley & Sons, Inc.

Vapnik V, Golowich SE, Smola A (1997) Support vector method for function approximation, regression estimation, and signal processing. Adv Neural Inf Process Syst 9:281–287.

Verzani J (2014) Package "gWidgets": gWidgets API for building toolkit-independent, interactive GUIs.

Walker D (2007) Similarity Determination and Case Retrieval in an Intelligent Decision Support System for Diabetes Management. Ohio University

White DJ, Thompson MJ, Vennapusa PKR (2007) Field Validation of Intelligent Compaction Monitoring Technology for Unbound Materials - Technical report for the Minnesota Department of Transportation (Mn/DOT). Minneapolis, USA

Williams GJ (2011) Data Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery, Use R! Springer

Willighagen E (2015) Package "genalg."

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

Yang J, Edwards DJ, Love PED (2003) A computational intelligent fuzzy model approach for excavator cycle time simulation. Autom Constr 12:725–735. doi: 10.1016/S0926-5805(03)00056-6

Zadeh LA (1965) Fuzzy Sets. Inf Control 8:338–353.

Zeiler M (1999) Modeling Our World: The ESRI Guide to Geodatabase Design. Redlands: ESRI Press

Zhang H (2008) Multi-objective simulation-optimization for earthmoving operations. Autom Constr 18:79–86. doi: 10.1016/j.autcon.2008.05.002

ANNEX A

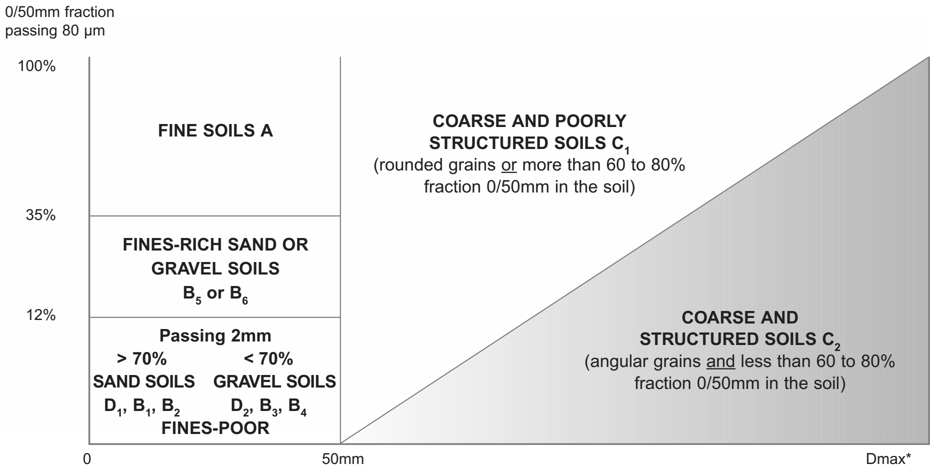


Figure A.1 Classification of soils according to grain size, where D_{max} refers to the maximum particle size (adapted from SETRA and LCPC 2000)

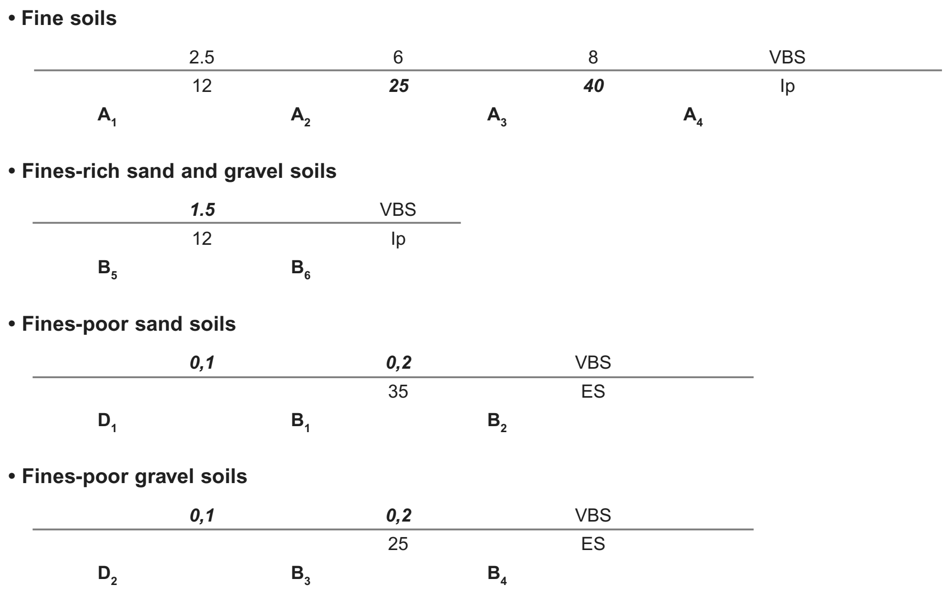


Figure A.2 Classification of soils according to both grain size and clay characteristics (adapted from SETRA and LCPC 2000)

Table A.1 Classification of soils according to their moisture state (adapted from SETRA and LCPC 2000)

Soil type	Reference test	State threshold			
		ts / s	s / m	m / h	h / th
A_1	IPI		25	***8***	***3***
	w_n/w_{OPN}	0.7	0.9	1.1	1.25
A_2	IPI		15	***5***	***2***
	w_n/w_{OPN}	0.7	0.9	1.1	1.3
	Ic	1.4	1.2	***1.05***	***0.9***
A_3	IPI		10	***3***	***1***
	w_n/w_{OPN}	0.7	0.9	1.2	1.4
	Ic	1.3	1.15	***1***	***0.8***
A_4	Special study required				
B_1		No sensivity to water content			
B_2	IPI			***8***	***4***
	w_n/w_{OPN}	0.5	0.9	1.1	1.25
B_3		No sensivity to water content			
B_4	IPI			***15***	***7***
	w_n/w_{OPN}	0.6	0.9	1.1	1.25
B_5	IPI		30	***12***	***5***
	w_n/w_{OPN}	0.6	0.9	1.1	1.25
B_6	IPI		25	***10***	***4***
	w_n/w_{OPN}	***0.7***	***0.9***	1.1	1.3
	Ic	1.3	1.2	1	0.8

ANNEX B

R_1 – Chalks – formed by the accumulation of falling calcite particles in the order of 1 to 10 μm in size. Their porosity contributes to the fragility of the material. Chalks are classified according to their dry density and moisture content (Table B.1);

Table B.1 Classification of chalks according to dry density and moisture content (adapted from SETRA and LCPC 2000)

$\rho d > 1.7$	R_{11}
$1.5 < \rho d \leq 1.7$ and $w_n \geq 27$	R_{12} h
$1.5 < \rho d \leq 1.7$ and $22 \leq w_n < 27$	R_{12} m
$1.5 < \rho d \leq 1.7$ and $18 \leq w_n < 22$	R_{12} s
$1.5 < \rho d \leq 1.7$ and $w_n < 18$	R_{12} ts
$\rho d \leq 1.5$ and $w_n \geq 31$	R_{13} th
$\rho d \leq 1.5$ and $26 \leq w_n < 31$	R_{13} h
$\rho d \leq 1.5$ and $21 \leq w_n < 26$	R_{13} m
$\rho d \leq 1.5$ and $16 \leq w_n < 21$	R_{13} s
$\rho d \leq 1.5$ and $w_n < 16$	R_{13} ts

R_2 – Calcareous rocks – This class contains the whole range of calcareous rock materials. Their predominant features, in respect of their use in fill, are their friability and, for the more fragmentable materials, frost susceptibility. The more compact calcareous rocks are classified according to their resistance in the micro-Deval test, while softer rocks are classified according to their bulk unit weight (Table B.2);

Table B.2 Classification of calcareous rocks according to their resistance in the micro-Deval test and bulk unit weight (adapted from SETRA and LCPC 2000)

$MDE \leq 45$	R_{21}
$MDE > 45$ and $\rho d > 1.8$	R_{22}
$\rho d \leq 1.8$	R_{23}

R_3 – Argillaceous rocks – These are characterised by a more or less resistant structure with a highly variable proportion of potentially swelling clay minerals imprisoned. These are classified mainly according to fragmentation and degradability tests. For the more fragmentable rocks, the natural moisture content is compared to either their normal Proctor optimum or their immediate bearing index, in order to determine their moisture state. (Table B.3);

Table B.3 Classification of argillaceous rocks according to their fragmentability and degradability (adapted from SETRA and LCPC 2000)

Fragmentability	Degradability	Class
$FR \leq 7$	$DG > 20$ $5 < DG \leq 20$ $DG \leq 5$	R_{31} R_{32} R_{33}
$FR > 7$	[$w_n \geq 1.3\ w_{OPN}$ or $IPI < 2$*] [$1.1\ w_{OPN} \leq w_n < 1.3\ w_{OPN}$ or $2 \leq IPI > 5$*] $0.9\ w_{OPN} \leq w_n < 1.1\ w_{OPN}$ $0.7\ w_{OPN} \leq w_n < 0.9\ w_{OPN}$	R_{34}th R_{34}h R_{34}m R_{34}s

R_4 – Siliceous rocks – This class of materials can be likened to assemblies of sand grains or stones cemented together with silica or calcite, where the strength of the binding affects the behaviour of the rock. The more compact rocks are classified according to their strength in the Los Angeles fragmentation test and micro-Deval wear test, while the softer rocks are classified according to their fragmentability (Table B.4);

Table B.4 Classification of siliceous rocks according to the Los Angeles and micro-Deval tests, as well as fragmentability (adapted from SETRA and LCPC 2000)

$LA \leq 45$ and $MDE \leq 45$	R_{41}
$LA > 45$ or $MDE > 45$ and $FR \leq 7$	R_{42}
$FR > 7$	R_{43}

R_5 – Saline rocks – In mechanical terms, this class of materials are like class R_2 and R_3 but they are more soluble in water and they are therefore liable to cause distress in the structure. They are classified depending on the proportion of salt and gypsum in the rock structure (Table B.5);

Table B.5 Classification of saline rocks according to their constitution (adapted from SETRA and LCPC 2000)

Soluble salt content (depending on degree of fragmentability): $\leq 5 - 10\%$ in rock salt $\leq 30 - 50\%$ in gypsum Slightly soluble salt rocks	R_{51}
Soluble salt content (depending on degree of fragmentability): $\leq 5 - 10\%$ in rock salt $\leq 30 - 50\%$ in gypsum Very soluble salt rocks	R_{52}

R_6 – Igneous and metamorphic rocks – This class of materials may have widely differing mechanical properties. Their fragmentability and friability may be very variable. The more compact rocks are classified according to their strength in the Los Angeles fragmentation test and the micro-Deval wear test, the softer rocks are classified according to their fragmentability (Table B.6).

Table B.6 Classification of igneous and metamorphic rocks according to the Lost Angeles and micro-Deval tests, as well as fragmentability (adapted from SETRA and LCPC 2000)

$LA \leq 45$ and $MDE \leq 45$	R_{61}
$LA > 45$ or $MDE > 45$ and $FR \leq 7$	R_{62}
$FR > 7$	R_{63}

ANNEX C

Table C.1 Classification for pneumatic tyred drum rollers (adapted from SETRA and LCPC 2000)

P1	CR between 25 and 40 kN
P2	CR between 40 and 60 kN
P3	CR greater than 60 kN

Table C.2 Classification for smooth vibrating drum rollers and vibrating tamping rollers (adapted from SETRA and LCPC 2000)

V1	$(M1)\times\sqrt{A0}$	between 15 and 25	and	$A0 \geq 0.6$
		greater than 25	and	$A0$ between 0.6 and 0.8
V2	$(M1)\times\sqrt{A0}$	between 25 and 40	and	$A0 \geq 0.8$
		greater than 40	and	$A0$ between 0.8 and 1.0
V3	$(M1)\times\sqrt{A0}$	between 40 and 55	and	$A0 \geq 1.0$
		greater than 55	and	$A0$ between 1.0 and 1.3
V4	$(M1)\times\sqrt{A0}$	between 55 and 70	and	$A0 \geq 1.3$
		greater than 70	and	$A0$ between 1.3 and 1.6
V5	$(M1)\times\sqrt{A0}$	greater than 70	and	$A0 \geq 1.6$

Table C.3 Classification for static tamping rollers (adapted from SETRA and LCPC 2000)

SP1	M1/L between 30 and 60 kg/cm
SP2	M1/L greater than 60 kg/cm but less than 90 kg/cm

Table C.4 Classification for vibrating plate compactors (adapted from SETRA and LCPC 2000)

PQ3	Mg/S between 10 and 15 kPa
PQ4	Mg/S greater than 15 kPa