



## EXPERT SYSTEM USING MULTI-OBJECTIVE OPTIMIZATION OF THE DIRECT CURRENT RAILWAY POWER SUPPLY SYSTEM

Manuel Soler Nicolau, Jesús López, Santiago Tapia, José Manuel Mera

*Research Centre on Railway Technologies (CITEF), Technical University of Madrid, Spain*

Submitted 13 July 2014; resubmitted 23 January 2015, 30 March 2015; accepted 13 May 2015;  
published online 02 November 2015

**Abstract.** There are many different aspects to be analyzed when designing a railway infrastructure. The energy system, which withstands the demand for energy from operating trains, must consider many factors to create a functional infrastructure, in terms of demanded energy and cost sustainable. The methodology proposed gives a set of possible solutions to the designer or engineer. On the one hand, this method works with a multi-objective genetic algorithm (NSGA-II), with high time efficiency. The main target of this work is to obtain the best electrical configuration in terms of number and location of substations and characteristics of the overhead line system. On the other hand, best configurations must take into account things such as real railway operation, signalling system, infrastructure, costs linked with environment, maintenance, construction and connection with general electric network, losses of energy dissipated along the catenary. Hence, this methodology must combine all of these skills and integrate it with a railway configuration, modelling and simulation tool, Hamlet developed at CITEF (Research Centre on Railway Technologies by Technical University of Madrid, Spain). After using this methodology, designers will have a set of configurations in order to get the final choice of location of traction substations and type of overhead line system to achieve properly the power demand from trains in railway systems.

**Keywords:** zone discretization; electric system optimization design; NSGA-II; maximum peak power demand; expert system.

### Introduction

Designing a system with multiple variables and conditioning factors, such as the implementation of the power supply system in a railway infrastructure, makes the task of the planner particularly difficult. Consideration has to be given to studies, both technical and budgetary, while assessing the impact it will have in environmental terms and, finally, having a future view to be aware of the difficulties, which might be posed by the selected design in operation and maintenance.

The thesis written by Johnston (1975) marks the interest in the study of optimization costs in railway infrastructures, and although in the years when this thesis was written there was still not complete mastery of electricity within railway energy sizing, the author was already considering the planning to obtain the costs of implementing the electrical design and, consequently, how to find a work or calculation tool to calibrate the cost-quality ratio of the design.

The leap in quality, occurs when methods are introduced, which leverage the studies reflecting the advan-

tage of a railway infrastructure depending on a certain criterion. These studies are also able to obtain a methodology, which improves those designs. The following studies specify the optimization techniques used as well as the impact they have on the ability to take decisions. Obviously, the progression in the computation capacity has had a clear effect on the simulation algorithms and tools, and clearly, the inclusion and adaptation of the computation capacities have gone hand in hand with the progress made in the formal optimization techniques.

There are several studies oriented to optimize the railway electrical grid, concerning in this task investment costs and reliability of the system. Related to this (Capuder *et al.* 2009), is working with an analytical study in order to improve the railway power system. To reach this purpose, the authors are using a simulator to get data and results and they basically are working in a DC system.

Working with genetic algorithm in the optimizing problem of supply system is Chuang *et al.* (2005). In this work is introduced the immune algorithm in the goal of



minimize the overall cost of power consumption and inverter investment in massive rapid transit. Another work with genetic algorithm is carried out in Pereira *et al.* (2014), searching the best place for rectifier substations.

On the other hand, there is PRT (Personal Rapid Transit), which power supply system allocation has different scope than in case of MRT (Metro Rapid Transit). The Viet *et al.* (2004), have developed an optimization process including load power balance and minimization of power loss. The technique integrated is an iterative approach.

Chang *et al.* (1999) have in mind the importance of optimizing the location of the traction substations. To do so, they use an algorithm of optimization by tabu search. Another of their contributions is to optimize energy consumption thanks to the firing angles necessary in current rectification, which is carried out specifically by the thyristors, which are part of the traction substations. With regard to the latter, in Chang *et al.* (1995), use is made of the optimization bi-criterion, a technique belonging to genetic algorithms. The advantage of an electrical system is evaluated depending on how the firing angles of the substations are programmed, making the optimum solution the one, which achieves the best compromise between uniform distribution between trains/substations and the power collected through the use of regenerative braking. Olofsson *et al.* (1995) is another interesting study related to optimization railway systems, working directly with the load flow in the line of study.

In the study carried out by Pilo *et al.* (2000), the main objectives of the authors, were to study the functional characteristics of an energy power system for high-speed railway systems, and thus a system of alternating current. The second objective they worked on was to optimize, or rather simplify, the electrical model by moving from a dual voltage model to models, which could be single voltage, thereby reducing the complexity of subsequent calculations (Jiménez-Octavio 2009), continuing with the route taken in the previous study, applies the same simulation model and modelling of the electrical system in such a way that, by implementing an optimization method based on genetic algorithms, more specifically multi-objective functions (Coello 2011), and contributing in some way to the MOGA algorithm (Murata, Ishibuchi 1995) to make the selection process more in tune with his requirements, obtains satisfactory results in terms of response times and the quality of the solutions provided.

Problems involved in multi-objective algorithms have not a unique optimum. The best solution to this problem depends on a number of factors before taking a decision, and normally these factors have opposing interests; in other words, if one is good, another undermines its result. In the end, a compromise must be reached between both. This problem has been considered in the same way, so that the more invested in costs of civil works, construction and installation, normally the smaller are the power losses in the network. Similarly, increasing the quality of the installation will result

in lower maintenance requirements. A multi-objective optimization algorithm would be perfectly adapted to this problem. The one chosen was the NSGA-II (Deb *et al.* 2002). A particular feature offered by this algorithm is the improved computational complexity it offers compared to its predecessors,  $O(MN^2)$ , where  $M$  is the number of objectives;  $N$  is the size of the population. Another factor in the choice of this algorithm is the simplicity in implementation wise, and thus of its application and adaptation.

To solve the decision-making process in the design of electrical systems, a sequential algorithm has been designed. For the different stages of this optimization strategy, a variety of software tools have been implemented, both for configuration, simulation, optimization and obtaining results. Tools were developed using the object-oriented language C++, and the multi-disciplinary mathematical tool *MatLab Release 2013a*. Given below are the keys to understanding the different phases of the complete process. The set of tools developed by CITEF (Soler *et al.* 2012; Mera *et al.* 2000), on which is enmarked Hamlet tool, is basic for tasks such as modelling, configuration, dynamic and electrical simulation of the railway system to be analyzed. The fundamental characteristics in each step of the algorithm are the integration with the genetic algorithm NSGA-II, the simplification proposed for the analysis of railway electric scenarios and using a complex system with multiple design variables and possibilities and thus solutions as a starting point. This method leads to a choice between various parameterized solutions, the value and analysis characteristics of which would already be known, hence making the problem much more accessible.

## 1. Method for Optimization of the Design of the DC Power System

Detailed below are the keys to understanding how results are obtained, which enable an analysis of a set of designs which offer improvements on initial electric system designs. The process is divided into three parts:

Creating the Discretization of the line in zones and weighting of the costs of the elements, such as substations and types of the overhead contact lines. In the same way, the discretized zones will be classified like maintaining, installing, environmental or easy to connect to general electric grid. Finally, a weight is assigned to each zone:

- applied artificial intelligence – characteristics of the genetic algorithm applied, NSGA-II;
- general algorithm – there is a detailed examination of the steps followed to obtain the best final solutions.

### 1.1. Discretization of the Railway Line. Weighting of Element Costs

In initial stage of designing the electrical power system, is very important to know the needs of the railway system. This involves the characteristics and linked budget of the different electric elements chosen, and the influ-

ence of their installation along the route. For this reason, in this study the final goal will be optimize the set of elements in order to avoid unnecessary money and impact in zones. The quantity of traction substations, the best position to inject sufficient power for the traffic demands and where they will be located are the skills of the methodology.

The core of the expert system considers discretizing the layout of the line, defining the critical zones and weighting the number of traction substations and zones.

*Discretizing the map. Zones of installation of traction substations*

The ability of the system depends on the action of discretizing the map belonging to the initial route of the circulations of the trains. During this initial step, the length of the line is analyzed. This length will be divided in zones, such as one traction substation will be able to be installed in each of this zone. In fact, this will be critical in the process of determining the positions of substations and if the participation of them is necessary in this zone or not.

In Fig. 1 there is an example. The route has been discretized in two zones, so this example has the possibility to install at most, two traction substations. In general, the discretization step must analyze the length of the route and depending on the initial number of traction substations planned to install, the line will be divided in  $n$  equal parts, where  $n$  is the number of substations. The length of these zones is constant. This ensures that the algorithm can evaluate all the positions of the railway layout and associate them at any time with a substation.

*Zones involved in evaluation*

In railway, electrical dimensioning design is mandatory take into account several impact factors, which they must be critical in terms of, evaluate budget and time in construction and maintaining the project. During the last years in CITEF, these critical factors have

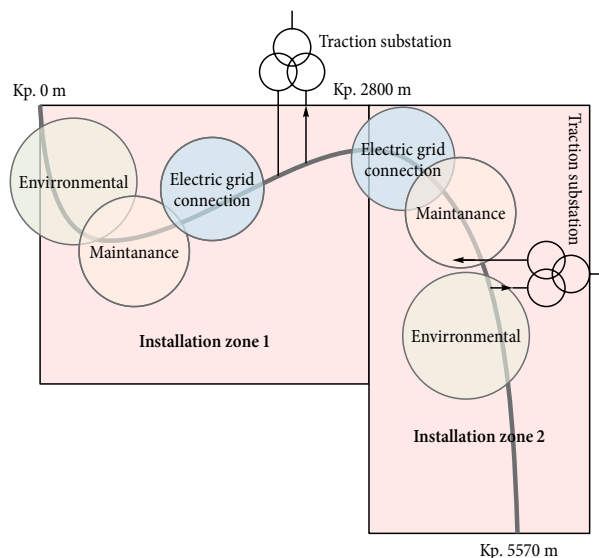


Fig. 1. Example configuration of discretization of the railway line

been linked with critical zones to study in this methodology:

- environmentally protected zones, which need to be avoided because of their high ecological value;
- zones for maintenance of the railway infrastructure; determine level of accessibility to carry out maintenance work;
- zones related to the evaluation of difficulty of connection between general electric grid and traction substation.

*Evaluation of components and zones*

Determining the cost or weight, in currency unit, of each element and zone involved in each project, is important to know if these costs are real or weighted. The methodology is capable to work with real or weighted cost but not mixed. If the designer/user of this methodology is able to collect the real cost of each element, all of them will be filled in the configuration files. However, if the option is weighting values, more flexible than the other, the process must take into account next issues:

- the relationship between the elements of cost must be provided or designed by the user;
- the substations, catenaries and discretized zones must have a value;
- if there is a zone of the line is not considered as zone maintenance, environmental, general network connection then it is not relevant in any case;
- high values penalize their adding – it is more difficult to enter as a component in the Pareto front with the final solutions.

**1.2. Applied Artificial Intelligence**

The algorithm NSGA-II has been implemented in C++, using the source code initially created in C (Deb 2006). This code has had to be adapted to this optimization methodology, to give it more flexibility for changes in configuration parameters of the genetic algorithm. Once migration was completed, the next step was to design and develop the intrinsic characteristics of the genetic algorithm in question:

Minimize objective functions:

$$y = f(x) = \{f_1(x), f_2(x)\} \tag{1}$$

subject to constraints:

$$\begin{cases} g(x) = \{g_1(x)\} \geq 0; \\ h(x) = \{h_1(x)\} = 0, \end{cases} \tag{2}$$

where:  $x = \{x_1, x_2, \dots, x_l\} \in X$  (decision space) or genotype.

*Genotype*

The genotype or individual, which corresponds to the set of variables in the problem, will be made up of  $n$  overhead line system (*ohl*) and  $m$  substations (*ss*). The number of overhead line system types, will be defined by the group of possibilities handled for the electrical design. Both variables are continuous. The  $m$  substations will be those initially planned.

$$\text{genotype} = \{ohl_1, ohl_2, \dots, ohl_n, ss_1, ss_2, \dots, ss_m\}, \quad (3)$$

where:

$$ohl \in [NegValue, 1];$$

$$ss \in [NegValue, 1],$$

where: *NegValue* is a low negative border to calculate the removing of substations.

For the substations, the method designed is such that, if the value of each of them is negative, the electrical algorithm will not take that substation into account, removing it from its installation zone. If the value is positive, the algorithm will have to modify the position of the substation according to that value. Knowing that each substation belongs to a known installation zone, the real positive value will indicate as a relative position, between [0, 1], the new position in which the substation should be located. This means that, if the zone has a length of 10km, and the value obtained is, for example, 0.5, the substation will be positioned at kilometre 5 of that zone. The Hamlet tool will receive the new position and will modify it within its configuration. *NegValue* is modifiable; if it is very negative, the probabilities to remove the substation increases; if it is very close to 0, the probabilities to remove a substation decreases, and the initial configuration is taken more into account.

Of all the overhead lines genes, the one with the highest value will be chosen, including this resulting overhead contact line in each installation zone. The electrical module of the Hamlet tool will receive information about the new overhead line system and will modify this information for the individual to evaluate.

*Construction costs objective function*  $f_1(x)$

The priority of this function is to obtain the total value of the cost involved in constructing the electrical system, which contains coding for each individual or genotype to be analyzed.

To reach this main aim, use will be made of the zonal discretization and the weighting or allocation of costs carried out in the previous steps. The allocation of costs is based on finding the intersection between the zones assigned to the installation of each traction substation and the other critical zones (environmental, connection to general electric network) before to the commissioning of the railway operation. It is also necessary to add the total cost of the elements included, the overhead line system chosen and the substations. Each genotype will configure its own scenario; hence, the location will depend of the substation to check if it will be constructed into a critical environmental zone or into a zone where the connection with the general electric grid is complicated.

$$CostCG_k = \delta_k \left( \frac{zie_k - zii_k}{zinst} \right) + \vartheta_k \left( \frac{zge_k - zgi_k}{zgen} \right); \quad (4)$$

$$CostCE_k = \delta_k \left( \frac{zie_k - zii_k}{zinst} \right) + \gamma_k \left( \frac{zee_k - zei_k}{zenv} \right). \quad (5)$$

$CostCG_k$  and  $CostCE_k$  compute the general cost related to the installation of the traction substation inside a general electric grid and environmental zone.  $zie_k$ ,  $zii_k$ ,  $zge_k$ ,  $zgi_k$ ,  $zee_k$ ,  $zei_k$  are the lengths belonging to the limits of the intersection between the installation of the traction substation and the respective zones, installation, general electric grid connection and environmental zones.  $zinst$ ,  $zgen$  and  $zenv$  are the lengths of the installation, general electric grid and environmental zones. The costs related to the zones are represented by  $\delta_k$ ,  $\vartheta_k$ ,  $\gamma_k$ .

Finally:

$$f_1(x) = \sum_{k=1}^{nsubst} (CostCG_k + CostCE_k + CostSs_k) + CostOhl, \quad (6)$$

where:  $nsubst$  is total number of substations;  $CostSs_k$  is cost of traction substation  $k$ ;  $CostOhl$  is cost of catenary.

*Objective function of power losses on the line and maintenance costs*  $f_2(x)$

In contrast to the above, this function attempts to evaluate the favourability of the design proposed by the individual or genotype, in relation to operating and maintenance costs.

The costs of maintenance zones, which have been established during the discretization and weighting steps, will be defined by the maintenance tasks, which must be carried out during the useful life of the railway infrastructure and which, in case of this method, will affect the overhead line system and electrical traction substations. It is based on the intersection between the zones of implementation of the substations and the zones established for the study of maintenance costs:

$$CostCM_v = \delta_v \left( \frac{zie_v - zii_v}{zinst} \right) + \rho_v \left( \frac{zme_v - zmi_v}{zmaint} \right), \quad (7)$$

where:  $CostCM_v$  is the total cost related to maintenance; the intersection between traction substation installation length and maintenance zone is determined by  $zme_v$  and  $zmi_v$ ;  $zmaint$  is the length of the maintenance zone;  $\rho_v$  is cost related to maintenance zone.

Another aspect, which must be minimized, is the power losses on the line. The calculation of power losses provides a measurement of quality of the chosen design.

To find the losses, data will be collected referring to the electrical demand of the trains in the cycles of maximum demand for electrical power; it will also be necessary to obtain the power values injected by the traction substations in those cycles:

$$Losses = \sum_{r=1}^{nsubst} Ssubst_r - \sum_{u=1}^{trains} Pet_u, \quad (8)$$

where:  $nsubst$  is number of traction substations in the energy system;  $trains$  is total trains circulating;  $Pet_u$  is electrical power demanded by train  $u$  in each cycle of time (the power factor assigned to the train, is applied in the calculus [VA]);  $Ssubst_r$  is power from substation  $r$  [VA].

Finally:

$$f_2(x) = \sum_{v=1}^{nsubst} (CostCM_v) + \alpha \cdot Losses, \quad (9)$$

where:  $\alpha$  is cost divided by VA (conversion factor).

Constraints

The restrictions, which have been implemented in the algorithm attempt to penalize heavily any situations, which cause the system to be non-viable.

The first restriction evaluates the convergence criterion, which will be explained in Section 2. It establishes that, if the calculation of the electrical scenario at the moment of maximum power demand, with the configuration proposed by the current genotype, does not fulfil this criterion, and consequently the substations cannot provide the necessary power to the trains, the restriction is not met. On the other hand, *conv* is an indicator that the scenario considered by the genotype would meet the requirements of the cycle with the greatest power demand. This function is a numerical method based in solving the general matrix, involving the power demanded by trains, the general impedance matrix and the searched voltages. In Section 2 is resumed, but the application in this methodology is carried out just by knowing if the system is electrically and energetically compatible with the operation and the dimensioning.

$$h(x) = conv(x) = \begin{cases} 1, & \text{if is not feasible;} \\ 0, & \text{if is feasible,} \end{cases} \quad (10)$$

where: *x* is an electrical scenario in a cycle (represented by genotype).

Next constraint is thought bearing in mind if the deviation of the costs of working with the dimensioning proposed is or not out of the ranges of the limits imposed by the engineers or the budgets studies. Hence there is a sum of the costs of the most relevant components and next step was deducted the budget limit and analyze if the difference is inside the tolerance values:

$$TotalCSc(x) = f_1(x) + \sum_{p=1}^{nsubst} CostCM_p. \quad (11)$$

Using last equation in the next constraint evaluation:

$$g(x) = budgetdev(x) = |ValConf - TotalCSc(x)|, \quad (12)$$

where: *ValConf* is an estimation by the engineer about the budget for the electrical scenario.

Once the optimization process is finished, the designer will have a Pareto Front. This Front contains the final best genotypes obtained by the process. Each of them are non-dominated by the others, that means the user must choose the selected electrical dimensioning design, represented by one genotype, bearing in mind the relationship values costs in both fitness functions, or evaluating better for the final purposes and objectives of the project, the best genotype in terms of any of both fitness values, independently of the value of the other. Finally the user will have its own decision-making step.

1.3. Characteristics of the General Algorithm

In all optimization of complex systems issues, such as the one being addressed, it is very important to reach

a compromise between the variables to be studied and evaluation functions together with the results obtained and, finally, the time taken to obtain them.

In the problem of optimization of the electrical design of railway infrastructures, the time impact of simulating full operation plans at the time of peak traffic, which makes it possible to analyze whether or not the electrical design is suitable, means that thought has to be given to limiting the complexity of the simulations. Fig. 2 shows the steps to follow in the designed optimization strategy, dividing steps belonging to Hamlet tool and steps specific to optimization method.

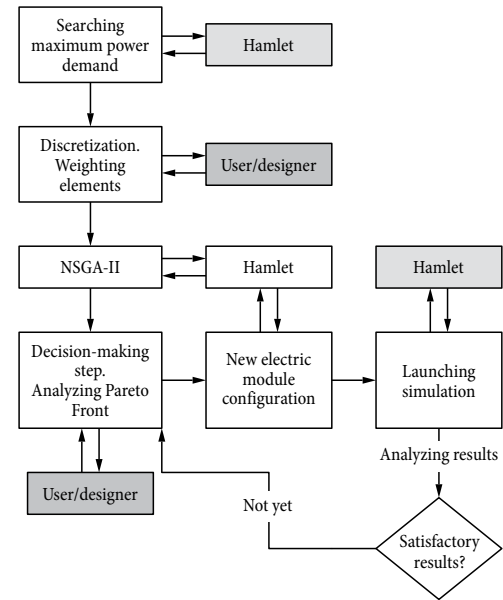


Fig. 2. Diagram optimization method

Hamlet tool steps

The main task of Hamlet tool in this strategy is to provide the data related to the electric power demanded by the trains during the operation, using real data about rolling stock, signalling, line data and electrics.

The first step corresponds to the modelling and configuration of the different modules (geometry, rolling stock, signalling and the initial electrical system) contained in the Hamlet tool. The initial electrical system can either be an oversized one, with more substations than necessary so that the algorithm can eliminate them, or it can use the requirements regarding railways installations considered in UNE-EN 50388:2006 CORR:2010 to design a more realistic first electrical scenario.

In the following step, it is worth highlighting the need to study the possibilities of traffic on the line. Because the purpose of the study is to achieve a selection of the overhead line systems and substations, and subsequently distribute them along the line to allow an power supply for the demand required by the trains at all times, the study must include an analysis of the situation with great stress for the line and, consequently, for the electrical framework. Consequently, the operational plans to be configured with the Hamlet tool must be those which, taking into account the capacities of the line, propose

greater traffic and a shorter interval between trains. The calculations and data for this stage must come from the study of passenger movement envisaged at rush-hour, the capacity of the rolling stock to handle the number of passengers, the signalling system, and add a percentage for analysis of traffic overload.

At next step, Hamlet is ready to launch simulation and get the results. It is considered a premise that, one of the critical aspects of study in the sizing of electrical railways is defined by the power demanded by the trains as they travel along the line. This electrical power calculated during simulation, would consist of:

– if traction:

$$S = \frac{F_m \cdot sp}{mechef} + Saux; \quad (13)$$

– if braking:

$$S = F_m \cdot sp \cdot mechef \cdot regcoeff + Saux, \quad (14)$$

where:  $S$  is the power obtained from the calculation of electrical traction forces [W];  $sp$  is speed of train [m/s];  $F_m$  is mechanical force by trains [N];  $mechef$  is mechanical performance, relation between mechanical and electric;  $Saux$  is power demanded by auxiliary engines in trains [W];  $regcoeff$  is regenerability coefficient due to brake.

If the demand is positive, the train is applying traction, and if it is negative, the train is braking; in this case the electrical module is ready to work on regenerative energy, and apply it to the calculation of the requirement of the train.

#### Optimization method steps

Once everything is configured, and Hamlet tool has collected each electric power demanded by trains, the next step is to achieve simplification of the problem so that the optimization problem is viable.

The intention is to find the moment of the simulation when the sum of all the  $F_m$  is maximum, which indicates the moment of greatest demand for the substations. This is the cycle, which needs to be found. Consequently, at this stage of the algorithm, with the figures for the power required by the trains for each cycle, the electrical scenario is put into action to check the energy stability of the system and then, with  $|S|$  (power) of each substation, all the values obtained from each substation are added together; this gives the estimated value for maximum power load on the line at that moment in time. When the simulation time is over, the figures for the moment with the greatest power demand will be stored and, consequently, if there is a study of the behaviour of different configurations of the electrical sizing with the conditions posed by this maximum peak, it is considered that the algorithm will provide solutions, which tolerate these load peaks.

The next step is to create the zonal discretization according to the planned initial electrical scenario. Then, the different variables belonging to the NSGA-II algorithm will be configured, such as the total of individuals, total of generations, parameters for selection, crossover and mutation. As explained in Section 1.2, the constraint

$h(x)$  requires us to know whether the electrical scenario is feasible or not. To do this, the figures for the cycle of maximum power demand are used: with these figures, the calculation of the electrical scenario is launched, just for that moment in time and with each genotype considered. The final result will say whether it is convergent or not, and whether the restriction is applied or not. One of the novel features of this work is precisely that. The simplification carried out allows a complex system to be optimized, examining a single moment in time, and thus ensuring that the duration of the algorithm is viable.

Finally, the multi-objective algorithm stores the results in files, from which the dominant Pareto Fronts are obtained. The last generation obtained will be the one, which offers the best results and can be used to analyze the railway infrastructure scenario; it is known that each of these will be convergent, from the previous process, thereby allowing a study of the results using other electrical parameters, such as minimum voltages in the overhead line, maximum currents or average quadratic currents in the substations. Those data can be collected using Hamlet and its electrical module. On each integration step time, will be stored electrical data along the simulation.

The final decision on the best distribution of substations and overhead line system selection will be taken by the engineer responsible for analyzing the data, as different solutions may be chosen depending on the response quality of the fitness functions provided by the algorithm.

## 2. Modelling Configuration and Simulation. Hamlet Tools

One of the objectives to be achieved by means of this tool is successful integration of analysis and optimization for various railway disciplines. Naturally, when one fully embarks on engineering projects as complex as railways systems are, the first factor to be taken into consideration is the need to combine and link up within one single tool the greatest possible volume of constituent sub-systems, in this case the analysis of railway lines. Based on this approach, the different modules integrated within Hamlet will be listed, along with the basic capacities and characteristics of each one.

As is described in Fig. 3, Hamlet has four main and interconnected parts.

### Infrastructure

Any simulator must have a main layer where the rest of the elements will be positioned. This is the main function of this module. Therefore, this is the supporting structure. In fact, there are a lot of possible configuration lines, so this infrastructure must to be ready to model every type of line and Hamlet has been useful for several configurations.

The basic configuration for the infrastructure would comprise sections and nodes. Each node would represent the location of points for track crossovers. It is important to point out that every signalling element is able to be installed using this planned infrastructure, elements such as signals, track circuits, balises, etc.



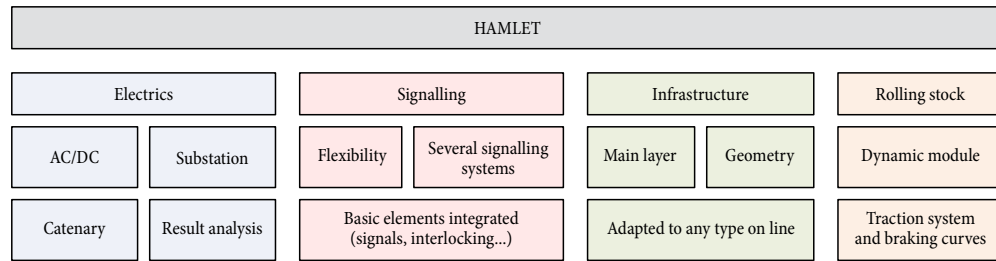


Fig.3. Modules and main skills of Hamlet

*Train*

For simulation of the rolling stock, Davis (1926) has been studied and employed, giving as the final function the balance between traction and resistive forces. Bearing in mind the motion function to be employed, which is general for any type of train, the other aspect, which must be generalized involves the configuration data, which must be common to any type of rolling stock, and thus allow for the simulation of a range of trains: metro, freight, local or high-speed.

*Electrics*

The electrical algorithm and model is devised in order to be able to simulate with direct current configuration, specifically for local/metro. The module is thus equipped to divide the line into the electrical sections into which the actual line is divided, and naturally, the electrical algorithm has integrated functions, which can include elements such as neutral zones, traction substations, overhead contact lines, etc.

A node-based analysis is performed in order to resolve the electrical system, with the different electrical elements making up the power supply system (substations, self-transformers, trains, static loads, etc.) being modelled by means of admittance matrices. The aim is to create a global admittance matrix grouping together the entire system, and taking into consideration all branches. The actual voltages and currents in the system are then calculated using an iterative process. The currents running along the resistive elements are defined by means of the power and the power factor being demanded. The process begins with the presumed voltages at all nodes making up the network, with these voltages, powers and power factors used to calculate the current source values:

- convergence: once the currents are known, the system is solved in order to obtain the new voltage values at the nodes, and these are compared with the previous values; if the difference between the two values is greater than a defined tolerance, the calculation is repeated, employing the voltages obtained in the previous calculation as the presumed values;
- not convergence: despite updating the voltages during the algorithm, not valid solution is found; that means the required power cannot be supplied with the designed system, and this scenario must not be included.

*Signalling*

Like the case of infrastructures, there are nowadays many signalling systems. Working with railway studies and simulations, Hamlet must be ready to model any of these types of signalling. Naturally, the tool is growing up, and the most important is that any new signalling system to develop in Hamlet, this tool is able to integrate it, because the main elements and the connection with the other modules do it an easy matter.

**3. Results**

The final effect of the tests, which have been carried out using this optimization strategy, is primarily to validate and check the methodology functions correctly. Consequently, the tests will be carried out with data obtained from a real project studied previously in CITEF. This project was chosen because the railway infrastructure adapts perfectly to the requirements specified with this optimization strategy, and on the other hand has interesting features for the study, such as for example a very pronounced height difference, which will increase traction efforts, and high regenerability effects in this area; this means that the variations in positioning and quantity of substations may be critical for the operation, which will be analyzed. The railway line is a metropolitan line 15 km long, with 13 stop stations.

To explain the tests, the path marked by the general algorithm is followed.

**3.1. Scenario configuration. Principal features**

- Geometry and elevated view: double track along the whole route and 13 stations. Speed restrictions have also been taken into account. Fig. 4 shows the profile, which must be transferred to the configuration, and also shows rapidly the height difference, which exists along the route.
- Rolling stock: a train formed of 6 cars, maximum speed 80 km/h, maximum service acceleration 1 m/s<sup>2</sup>, maximum service deceleration 0.9 m/s<sup>2</sup>. These would be the most important data to be detailed, although the rolling stock is configured with all the data relating to modelling, which was introduced in Section 2.
- Signalling: The planned signalling system for this line was CBTC (IEEE Std 1474.1-2004). In the corresponding Hamlet module, the basic characteristics of this system were implemented:
  - moving blocks;

- minimum safety distance between trains (30 metres).
- Electrical scenario: The line is Direct Current at 1500 V. Two types of overhead line systems (Table 1) have been chosen. The configuration tool is able to configure with the set of conductors, but here is shown only the main characteristics.

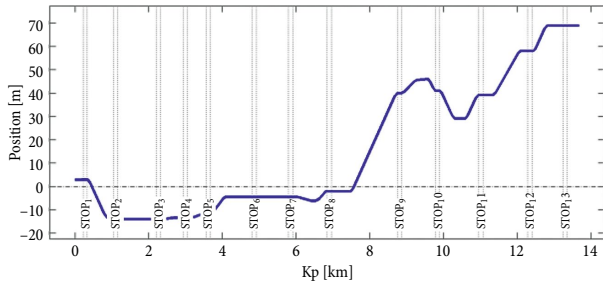


Fig. 4. Elevation of the line

For the initial scenario, 7 electrical traction substations are installed to begin with, as specified in the real project. In our design, that means a genotype of seven bits (continuous) representing the traction substations and two bits (continuous) representing the catenaries. Initial configuration is shown in Table 2.

Table 1. Main characteristics of overhead contact lines

	Section [mm <sup>2</sup> ]	Resistance [Ω/km]	Resistivity [Ω·mm <sup>2</sup> /m]
OHL <sub>1</sub>	1080	0.01897	0.02049
OHL <sub>2</sub>	1277	0.01487	0.01899

Table 2. Initial configuration of electric dimensioning

SS <sub>1</sub>	SS <sub>2</sub>	SS <sub>3</sub>	SS <sub>4</sub>	SS <sub>5</sub>	SS <sub>6</sub>	SS <sub>7</sub>	Type of catenary
1.5	2.3	5.8	7.3	9	10.7	13.5	OHL <sub>1</sub>
Kp	Kp	Kp	Kp	Kp	Kp	Kp	installed

The next step, according to the algorithm, is to define the operating plan, which can be considered critical for the stress analysis of the electrical sizing. According to the real project conditions, the following characteristics are established – trains with a 90 headway. Stopping times in stations are 15–20 seconds, depending on the station stop, which results in a fleet of 80 trains

Finally, the main characteristics of the machine used for this optimization process are I5-2400 3.1 GHz 4GB RAM.

### 3.2. Zonal Discretization and Weighting of Costs

In the experiments carried out to validate the method, it is not necessary to obtain real data of costs and zonal distribution. What has been sought is to give the experiment sufficiently broad diversity so that test mechanisms can be established on the point being dealt with, and so that the genetic algorithm has sufficient working evaluation characteristics. The following zonal diversity and allocation of costs has been established in Tables 3–4. This

discretization and weighting has been designed without real data and the goal is taking into account one possible configuration, on which the set of types of zones have representation.

Once the configuration stage is completed, the mechanical and electrical simulation is started. When initial simulation data is collected and final optimization results are collected, is able to compare both results and get conclusions about the optimization design proposed by the algorithm.

The graphs of minimum voltages in the catenary are a very useful indicator for reaching conclusions about the railway operation: there can be quick analysis of the areas of greatest demand, which can be located on the railway line. It is also possible to see whether any area is critical in terms of energy demand.

The graph of minimum voltages for the original scenario, Fig. 5 shows major rises and falls in voltage around the area of greatest height difference, but it may also be concluded that around some substations the voltage provided is not very high, making it very probable that the scenario is oversized, as will be seen in the optimized solutions. It may be seen, that the voltages never reach the rated voltage (1500 V). This is due to the fact that we are working with real models, taking

Table 3. Substations and overhead contact line weight

Element	OHL <sub>1</sub>	OHL <sub>2</sub>	SS <sub>1</sub>	SS <sub>2</sub>	SS <sub>3</sub>	SS <sub>4</sub>	SS <sub>5</sub>	SS <sub>6</sub>	SS <sub>7</sub>
Weight	300	400	100	100	100	400	100	100	400

Table 4. Discretization zone and weight

Element	Type	Initial position [m]	Final position [m]	Weight
Zone 1		0	2000	135
Zone 2		2000	4000	50
Zone 3		4000	6000	90
Zone 4	Installation	6000	8000	50
Zone 5		8000	10000	135
Zone 6		10000	12000	50
Zone 7		12000	13600	50
Zone 8		2000	3000	100
Zone 9	Environmental	9500	10500	100
Zone 10		4000	4250	3000
Zone 11		100	750	1000
Zone 12		5000	7700	200
Zone 13	Connection to general grid	500	1700	250
Zone 14		3000	3500	300
Zone 15		8000	9200	1500
Zone 16		0	1000	650
Zone 17		1000	3550	650
Zone 18	Maintenance	4200	4500	500
Zone 19		6000	7700	200
Zone 20		8000	8650	3350



into account the losses produced by the corresponding impedances. A check is made to ensure that, according to standard UNE-EN 50329:2004/A1:2011, the values are within the range {1000 V, 1800 V}.

### 3.3. Analysis of Results of the Optimization

As proposed in the optimization strategy developed, the data provided by the simulation of the complete scenario with the initial configuration and the analyzes of the results, the objective is to find the cycle of maximum power demand, so that it can be used in the calculations of convergence, which will be needed by the different individuals generated by the algorithm NSGA-II.

The most significant values applied to the genetic algorithm for the optimizations, which will be described below, are those appearing in Table 5. These final values fulfil the skills demanded by this optimization method. Last generation has enough individuals to check the goodness of each configuration defined by them. Finally the results were obtained with a population of 10/20 individuals. Both results were checked and analyzed.

Once the genetic algorithm has been executed, a population is obtained in the last generation with dominant values for the two objective functions considered. Presented below are the electric results deriving from configuring the scenario using latest generation results, and selecting those individuals, which provide the minimum for each of the functions.

#### Genotype with best fitness value for objective function relative to operational/construction costs

Fig. 6 shows the data for minimum voltages in the pantograph, for the scenario proposed by the latest generation genotype, which had the best fitness for the objective function belonging to the construction costs (1594). The solution comprises the  $OHL_2$ , and the substations 1, 2, 4 and 7, and the remainder are removed. The decisions taken to choose one or another solution depend at all times on the human factor. For example, in this case, although it can be seen from Table 6 that the viability analyzes, studied in this case because of the short and long duration overloads, are within acceptable parameters according to European regulations UNE-EN 50329:2004/A1:2011, other considerations need to be taken into account, such as studies of the system in degraded situations such as, for example, a fault in a substation set, or directly in the whole substation. Furthermore, another factor, which must be included, and which stands out in Fig. 6, this is the cost value of operation (losses + maintenance). This value is considerably higher than the average, and consequently it is possible that, depending on factors such as useful life of elements, infrastructure, or simply because of budget constraint, this solution may be unviable. Finally, the conductors of the overhead line system chosen, have a larger cable section and lower resistivity, and are also more costly than the alternative.

#### Genotype with best fitness value for objective function relative to exploitation costs

This solution, logically, is less stressful for the electrical system, as can be examined in Fig. 7, than the previous Fig. 6. From the analytical point of view, where in the original scenario we had substations 5 and 6, for a better solution for the losses plus maintenance function, these substations are not necessary; however, unlike the previous test example, substation 3 is included, and is also new, installing the overhead line system  $OHL_2$ . Obviously, an immediate consequence is that the losses will be much smaller. A pattern related to the previous test is that the best option for values obtained by this function, involves high construction costs, as can be seen in the Fig. 8.

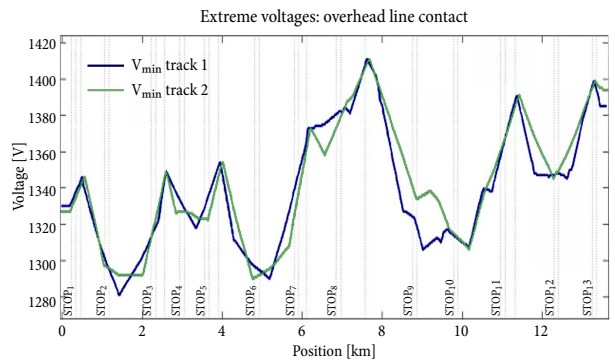


Fig. 5. Minimum voltage results in initial configuration

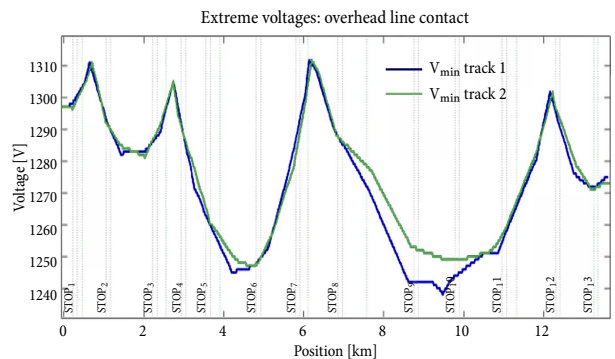


Fig. 6. Minimum voltage in case 1

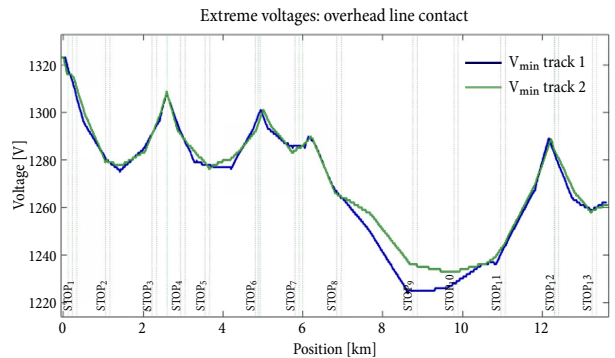


Fig. 7. Minimum voltage in case 2

Table 5. NSGA-II main characteristics

Population size	Generations	Crossing probability (0–1)	Mutation probability (0–1)	Distribution index mutation (0–100)	Distribution index cross-over (0–100)
10/20 individuals	25	0.7	0.3	350	25

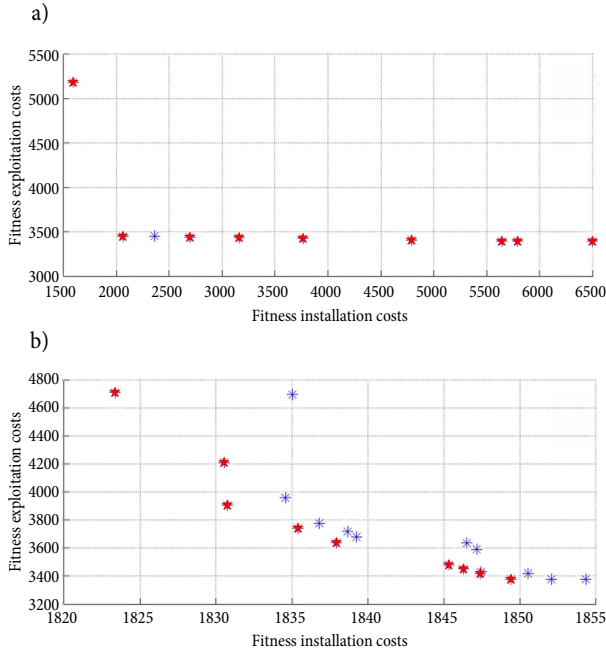


Fig. 8. Pareto front: a – 10 individuals; b – 20 individuals

For a complete viability study, as has been suggested previously, there should be a study of the possible degraded situations: as can be observed in Table 6, the values for overloads, minimum voltages and maximum currents in the contact wire are within the permitted margins.

### 3.4. Comparison

The analysis of the short and long duration overload was carried out by calculating  $I_n$  (rated current) in the substations, using Hamlet simulator. For this project and the level of traffic studied,  $2 \times 3500$  kV transformer sets were planned for each substation, making the rated power ( $S_n$ ) 7000 kVA. Given that the rated voltage ( $V_n$ ), is known, 1500 V in the case of this operation, we obtain the following:

$$I_n = \frac{S_n}{V_n} = \frac{7000}{1500} = 4666.66 \text{ A.} \quad (13)$$

It can be seen, that the minimum voltage and maximum current values in the contact wire are a little lower and higher, respectively, although they are clearly within the permitted margins described in the normative. As far as the RMSs (Root Median Square) are concerned, the increase is also considerable, due to the substations, which are not implemented in the original locations.

In Fig. 8a shows how the algorithm works with a relative low population. The final generation gives a front practically without discarded individuals, and there are a great set with similar exploitation cost fitness. Nevertheless in Fig. 8b, using 20 individuals in the configuration, the final sample is more heterogeneous. Clearly show different levels of fronts and the Pareto front has more values to analyze in order to reach the best scenario represented for each one.

Finally, a brief about the computational times is shown in Table 7.

Table 6. Main electric comparative results

	Initial configuration	Best fitness $f_1$	Best fitness $f_2$
$V_{min}$ track 1 [V]	1281	1238	1224
$V_{min}$ track 2 [V]	1290	1247	1233
$I_{max}$ track 1 [A]	1510	1875	1890
$I_{max}$ track 2 [A]	1257	1856	1686
SS <sub>1</sub> RMC (1 min) [A]	4412.3	5809.7	5386.6
SS <sub>1</sub> RMC (30 min) [A]	3640.5	4819.0	4473.5
SS <sub>2</sub> RMC (1 min) [A]	4186.7	6269.6	1.7
SS <sub>2</sub> RMC (30 min) [A]	3452.3	5184.6	1.4
SS <sub>3</sub> RMC (1 min) [A]	2994.6	Not included	6240.3
SS <sub>3</sub> RMC (30 min) [A]	2445.1		5044.4
SS <sub>4</sub> RMC (1 min) [A]	2985.2	5845.8	6056.5
SS <sub>4</sub> RMC (30 min) [A]	2472.8	4752.8	5022.4
SS <sub>5</sub> RMC (1 min) [A]	2649.9	Not included	Not included
SS <sub>5</sub> RMC (30 min) [A]	2237.0		
SS <sub>6</sub> RMC (1 min) [A]	4862.5	Not included	Not included
SS <sub>6</sub> RMC (30 min) [A]	4032.7		
SS <sub>7</sub> RMC (1 min) [A]	4566.9	5164.9	5348.9
SS <sub>7</sub> RMC (30 min) [A]	3814.4	6227.1	6480.2

Table 7. Computational times of the optimization process

Situation	Time [s]
Average time per genotype	0.27
Solving electrical equations	0.11

The optimization process in this test takes between one and two hours, with a population between 10/20 individuals, respectively. It is important to point out that the length of the whole optimization is determined by the high computational time due to the electrical solver.

### Conclusions

The decisions taken for the design of an energy system for railway infrastructures involve the application of a variety of modelling, calculation and simulation tools, to finally obtain data allowing us to evaluate the alternatives. The scenarios presented in this article show, that for a direct current railway line, with a selection range of 2 overhead line systems types and an initial design of

7 substations, the algorithm repositions them according to the costs and power losses on the line and chooses the best overhead line system to distribute the current to the trains operating on the line. In a railway infrastructure it is important to know and manage the environmentally protected areas, areas with a difficult connection to the electric general network and critical areas for maintenance works. Zonal discretization provides an effective method of solving this problem.

The flexibility of associating costs to the model is another functionality, which brings the best solution closer to the person who finally has to decide on the best possible design. Furthermore, the operation can be analyzed from practically all the important aspects of a railway structure, thanks to the modelling of infrastructure, trains, electricity and also the signalling system.

Taking into account the method of searching the maximum power demanded peak during the planned operation that has reached a very important improvement. Because the methodology keeps the power of the electric calculation algorithms, which are able to analyze the convergence of the proposed energy system and this method is aware about the computational time. Instead of analyzing the whole integrations steps belonging to the operation, the method just analyze one. Bearing in mind an operation of 2 hours and a step of integration of 0.5 s, that would result  $0.5 \cdot 7200 = 3600$  electric convergence calculations. Using this strategy, one analysis is done to get an approximation of the validity of the electric system designed. Thinking in terms of time, with the example related above, there would have been necessary between 250–500 design analyzes. This strategy transforms this kind of studies in an optimization acceptable problem.

This area of research presents different directions for future work, which can be classified into four main ones. One is to extend the range of work to various types of energy sizing, such as high-speed lines with alternating current, which will include autotransformers, neutral zones, etc. Related to this point, the evaluation of the costs by the objective functions can include time use variables, as well as stochastic parameters. An integrated analysis of degraded electrical situations, combined this in convergence criterion and constraints, is an important point to develop.

## References

- Capuder, T.; Lugaric, L.; Brekalo-Strbic, J.; Krajcar, S. 2009. Optimizing the train power system in Zagreb, in *VPPC'09: IEEE Vehicle Power and Propulsion Conference*, 7–10 September 2009, Dearborn, MI, 41–45. <http://dx.doi.org/10.1109/VPPC.2009.5289872>
- Chang, C. S.; Low, J. S.; Srinivasan, D. 1999. Application of tabu search in optimal system design and operation of MRT power supply systems, *IEE Proceedings – Electric Power Applications* 146(1): 75–80. <http://dx.doi.org/10.1049/ip-epa:19990214>
- Chang, C. S.; Wang, W.; Liew, A. C.; Wen, F. S.; Srinivasan, D. 1995. Genetic algorithm based bicriterion optimisation for traction substations in DC railway system, in *IEEE International Conference on Evolutionary Computation 1995*, 29 November – 1 December 1995, Perth, WA, Australia. <http://dx.doi.org/10.1109/ICEC.1995.489111>
- Chuang, H. J. 2005. Optimisation of inverter placement for mass rapid transit systems by immune algorithm, *IEE Proceedings – Electric Power Applications* 152(1): 61–71. <http://dx.doi.org/10.1049/ip-epa:20041143>
- Coello, C. A. C. 2011. Evolutionary multiobjective optimization, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 1(5): 444–447. <http://dx.doi.org/10.1002/widm.43>
- Davis, W. 1926. *The Tractive Resistance of Electric Locomotives and Cars*. General Electric.
- Deb, K. 2006. *Original Implementation (for Windows and Linux): NSGA-II in C (Real + Binary + Constraint Handling)*. Kanpur Genetic Algorithms Laboratory, India. Available from Internet: <http://www.iitk.ac.in/kangal/codes.shtml>
- Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation* 6(2): 182–197. <http://dx.doi.org/10.1109/4235.996017>
- IEEE Std 1474.1-2004. *IEEE Standard for Communications-Based Train Control (CBTC) Performance and Functional Requirements*. <http://dx.doi.org/10.1109/IEEESTD.2004.95746>
- Jiménez-Octavio, J. 2009. *Análisis dinámico y optimización de catenarias para alta velocidad*: Tesis Doctoral. Universidad Pontificia Comillas, Madrid, España. (in Spanish).
- Johnston, W. D. 1975. *Mainline Railway Electrification: an Economic Feasibility Model*: MSc Thesis. University of British Columbia. Canada. 194 p. Available from Internet: <https://circle.ubc.ca/handle/2429/19118>
- Mera, J. M.; Tapia, S.; Vera, C.; Jaen, J. A. 2000. Railway lines operation simulator: GifTren, in *Seventh International Conference on Computers in Railways: Computers in Railways VII*, 11–13 September 2000, Bologna, Italy, 997–1006.
- Murata, T.; Ishibuchi, H. 1995. MOGA: multi-objective genetic algorithms, in *IEEE International Conference on Evolutionary Computation 1995*, 29 November – 1 December 1995, Perth, WA, Australia. <http://dx.doi.org/10.1109/ICEC.1995.489161>
- Olofsson, M.; Andersson, G.; Söder, L. 1995. Optimal operation of the Swedish railway electrical system, in *International Conference on Electric Railways in a United Europe 1995*, 27–30 March 1995, Amsterdam, Netherlands, 64–68. <http://dx.doi.org/10.1049/cp:19950179>
- Pereira, F. H.; Lobo Pires, C.; Ikuyo Nabeta, S. 2014. Optimal placement of rectifier substations on DC traction systems, *IET Electrical Systems in Transportation* 4(3): 62–69. <http://dx.doi.org/10.1049/iet-est.2010.0063>
- Pilo, E.; Rouco, R.; Fernández, A.; Hernández-Velilla, A. 2000. A simulation tool for the design of the electrical supply system of high-speed railway lines, in *IEEE Power Engineering Society Summer Meeting 2000*, 16–20 July 2000, Seattle, WA, 2: 1053–1058. <http://dx.doi.org/10.1109/PESS.2000.867519>
- Soler, M.; López, J.; Mera, J. M. 2012. Simulation system for the optimization of a block distribution under the ERTMS-1 signalling system, in *Thirteenth International Conference on Design and Operation in Railway Engineering (COMPRAIL 2012)*, 11–13 September 2012, New Forest, UK, 61–72.
- UNE-EN 50388:2006 CORR:2010. *Aplicaciones ferroviarias. Alimentación eléctrica y material rodante. Criterios técnicos para la coordinación entre sistemas de alimentación (subes-*

tación) y el material rodante para alcanzar la interoperabilidad [Railway Applications. Power Supply and Rolling Stock. Technical Criteria for the Coordination Between Power Supply (Substation) and Rolling Stock to Achieve Interoperability] (in Spanish).

UNE-EN 50329:2004/A1:2011. *Aplicaciones ferroviarias. Instalaciones fijas. Transformadores de tracción* [Railway Applications. Fixed Installations. Traction Transformers] (in Spanish).

Viet, N. X. H.; Song, H.-S.; Nam, K. 2004. Locating power supplies on a personal rapid transit system to minimize system losses, *IEEE Transactions on Industry Applications* 40(6): 1671–1677. <http://dx.doi.org/10.1109/TIA.2004.836310>

## APPENDIX

$CostCG_k$  is cost related to the installation of substation  $k$  in a general electric grid zone;

$CostCE_k$  is cost related to the installation of substation  $k$  in a environmental zone;

$CostCM_v$  is cost related to the installation of substation  $v$  in a maintenance zone;

$CostOhl$  is cost related to catenary;

$CostSs_k$  is cost related to traction substation  $k$ ;

$Fm$  is mechanical force by trains [N];

$mechef$  is mechanical performance, relation between mechanical and electric [0,1];

$ohl$  is a bit belonging to genotype, which represents catenary;

$Pet_u$  is electrical power demanded by train  $u$  in a cycle of time; the power factor assigned to the train, is applied in the calculus [VA];

$regcoeff$  is regenerability coefficient due to brake [0,1];

$S$  is the power obtained from the calculation of electrical traction forces [W];

$Saux$  is power demanded by auxiliary engines in trains [W];

$sp$  is speed of train [m/s];

$ss$  is a bit belonging to genotype which represents traction substation;

$Ssubst_r$  is power from substation  $r$  [VA].

$ValConf$  is a preconfigured value, which represents the maximum value allowed for any scenario;

$zee_k$  is end position of the substation  $k$  inside environmental zone [m];

$zei_k$  is initial position of the substation  $k$  inside environmental zone [m];

$zenv$  is total length belonging of a environmental zone [m];

$zgen$  is total length belonging of a general electric grid zone [m];

$zge_k$  is end position of the substation  $k$  inside general electric grid zone [m];

$zgi_k$  is initial position of the substation  $k$  inside general electric grid zone [m];

$zie_k$  is end position of the substation  $k$  inside installation zone [m];

$zii_k$  is initial position of the substation  $k$  inside installation zone [m];

$zinst$  is total length belonging of a installation zone [m];

$zmaint$  is total length belonging of a maintenance zone [m];

$zme_v$  is end position of the substation  $v$  inside maintenance zone [m];

$zmi_v$  is initial position of the substation  $v$  inside maintenance zone [m];

$\alpha$  is a conversion factor, representing cost divided by VA;

$\Theta$  is cost related to a general electric grid zone;

$\delta$  is cost related to a installation zone;

$\gamma$  is cost related to a environmental zone;

$\rho$  is cost related to maintenance zone.