An optimisation model to locate level crossings in railway lines at mines to minimise the total weighted-walked distance

Abstract

The location of level crossings for pedestrian circulation is a safety issue due to potential human losses and material damage. This problem has become a concern in urban and industrial environments due to the characteristics of rail systems and unsafe pedestrian behaviour. Research reveals that psychological and infrastructural factors influence pedestrians' decision-making when crossing railroad lines. One of these is the convenience that leads pedestrians to cross at the nearest point. However, investigations to date only study the impact of location on the risk and severity of crashes and pedestrian rule violations. The current research is quasi-experimental, i.e., no optimisation tool has been developed to determine which crossing points pedestrians should follow. Therefore, this research develops a mathematical model for locating level crossings at railways to minimise the total weighted-walked distance. A genetic algorithm was proposed for solving the model, especially for large-size problems. The algorithm parameters were calibrated using the design of experiments and ten instances based on the characteristics of the area under study. The obtained results provide (1) the location of the level crossings and (2) the route pedestrians should follow for each origin-destination pair based on the facilities adjacent to the rail lines. A sensitivity analysis was performed using the ten instances to determine how the location of the level crossings changes with modifications in the facilities' location. Results showed that the model provides alternative solutions according to the problem's size. This study delivers a methodology for practitioners and stakeholders of transportation systems.

Keywords: Level crossing, Pedestrian, Safety, Weighted distance, Modelling

1. Introduction and background

Accidents at rail level crossings (RLXs) between trains and other system users are one of the most critical problems in railway areas since they trigger human and material losses (such as injuries, deaths, insurance payments, and legal fees) that jeopardise safety on these tracks [1]. Moreover, collisions at RLXs are the leading cause of fatalities in railways. Thus, safety in these systems has become a worldwide issue [2]. According to the Federal Railroad Administration [3], 4944 accidents occurred in the United States between 2019 and 2021, where 36% of them resulted in fatalities or injured people. In the European Union, 1484 rail accidents led to death, while 1080 recorded injuries in 2020 [4]. In addition, 31% and 37% of those accidents involved RLX users. However, local/national government studies concluded that train-pedestrian accidents often involve trespassers, "people who are illegally on the rail" [2].

This problem is found in urban environments and industrial sectors such as mining. Rail transport has become a key element in mining operations because of its effectiveness and economy [5]. For example, the mining industry in Australia invests approximately US$1.5 billion in rail transport [6]. In Colombia, 15% of cargo (mainly coal) is moved by rail [7]. However, the mining industry still faces safety challenges due to the characteristics of its railway system, as follows [8]:

- Railways at mining often have more than one line, so it is common to find trains travelling in both directions.
- Trains do not travel on fixed schedules because it depends on the mining operation.
- It is difficult to predict the speed of an oncoming train.
- Modern trains are quieter.
- Trains always have the right of the way and take a long time to stop.
- Train drivers' visibility is quite limited, i.e., they cannot see people located in front of the train.
- The probability of surviving a train crash is almost zero due to its weight and inertia.

Authorities and stakeholders of the railway industry have promoted strategies to reduce the number of fatalities, mostly at public RLXs. On their side, researchers have proposed different interventions to reduce train-pedestrian accidents. As a result, multiple findings have tried to provide logical explanations for the issues involved in RLXs within cities or on roads connecting towns and cities. Many factors lead pedestrians to have unsafe behaviours, where the main one is convenience (less time and effort) because they tend to choose the shortest path [9, 10]. However, other variables and risk factors influence pedestrians' decisions and the crashes' frequency, severity, and occurrence. Results conclude that the risk factors can be summarised as pedestrian/drive gender [11, 12, 13], frequency of RLX use [11], sensation seeking [11], perceived hazard [14], controls/warning devices [14, 15], train position [14], descriptive norms [14], type of vehicle/train [12], period of the day [16, 12], weather [12], age [12], and the number of rail tracks [17].

Literature has also shown that the RLX's infrastructure impacts the accident severity between trains and pedestrians. Active crossings are equipped with automatic control to help users know when to cross. On the other hand, passive crossings have "stop" signs, and the user crosses if it is safe. Most studies agree that RLXs with active controls are much more effective in avoiding violations than passive ones because drivers' obedience is higher at active RLXs [1, 18] . Also, flashing lights are more effective than traffic lights in warning drivers, especially the red ones [19]. Other authors have looked at psychological risk factors that justify the complexity of RLXs. Read et al. [20] concluded that decision-making at RLXs is not as straightforward as it may appear at first glance, i.e., many options, strategies, and influencing factors must be considered. Salmon et al. [21] highlighted the driver's mental schema as the main cause of the accident between a passenger train and a loaded semi-trailer. Beanland et al. [22] determined that familiarity with the RLX is a risk factor that generates a difficult-to-change mental schema in which drivers expect to find a familiar situation. Another issue that explains RLX violations is mistakes made due to users' lack of knowledge of the zone's rules [22, 23].

Recently, Read et al. [24] conducted a literature review on systematic and problematic factors at RLXs. They used AcciMap to consolidate all influencing elements and classify them using several categories. One of them was *equipment and surroundings*, and findings determined that the *location of the RLXs* is related to unsafe behaviours in pedestrians. However, most studies tend to understand users' behaviour, psychological characteristics of users and the physical attributes of the RLX itself. Few works have considered the location of RLXs as a factor of accident severity, unsafe behaviour, or risk perception. Anandarao and Martland [25] applied probabilistic risk assessment techniques on six RLXs in Japan, considering several influential factors. The exploratory analysis concluded that the location is an attribute of the RLXs that affects the crossing risk. Moreover, RLXs with less than 20 meters of visibility have a 50% higher accident rate than those with less distance. Davey et al.

[26] collected the experiences of vehicle and train drivers at RLXs to identify the factors related to the incidents and accidents. Results yield that the crossing location is one of the main contributors to increasing the risk. In addition, the design and location of RLXs affect the drivers' visibility. Palat et al. [27] conducted several surveys to identify the factors that explain violations at RLXs in pedestrians and car drivers in the suburbs of Paris. The authors found that the RLX location was significant in developing risky crossing behaviours because of the environment surrounding the RLX, such as the traffic and nearby stations. Liang et al. [28] analysed the violation rate of motorists considering the RLX location (railway station nearby or not), closure cycle, and road traffic density. They concluded that the violations are due to variability in the timing during the last phase of the closure cycle (when the barriers are down). However, the latter is related to the location of the RLXs, i.e., the closer the RLXs are, the higher the violation rate.

As seen from the above studies, the location of RLXs determines risk, accident and pedestrian violation rates. It means that RLXs should not be located empirically but based on robust decision-making tools to optimise indicators, evaluate scenarios and adapt current solutions to changes in the system. However, existing studies are limited to quasi-experimental or descriptive research. It means the literature fails to provide research work that proposes a quantitative methodology based on optimisation techniques to determine the location of RLXs in railways. Moreover, no research of this type has been performed on the mining sector, where the system's characteristics make the problem more complex. Therefore, the novel contribution of this paper is to develop a mathematical model for locating RLXs in a railway system and a genetic algorithm for its solution. This model considers the pedestrian generation points (origins and destinations) and the location of the rail lines. In addition, a sensitivity analysis will be performed to determine how the solution calculated by the model changes with modifications in the system and the flexibility of the proposed model. This study strengthens the literature by offering a methodology based on optimisation that supports the practitioners' decision-making.

The rest of the paper is divided as follows. Section 2 describes the problem and develops the mathematical formulation. The solution methodology is presented in Section 3 while Section 4 displayed the computational results and discusses the main findings. Finally, Section 5 concludes the final remarks and suggests future works.

2. Problem description and formulation

The study is focused on the Dispatch Centre of a railway system within an open-pit mine in northern Colombia. This area spans 3220 square meters and accommodates four railway lines, each capable of transporting 109 wagons. The railway connecting the mine to the maritime shipping terminal covers a distance of 150 kilometres. Due to the diverse administrative and maintenance facilities on either side of these rail lines, personnel frequently need to cross them due to the nature of their roles. In addition, individuals must transit this area to access the maintenance shop for either heavy or light equipment. Consequently, it becomes imperative to determine the optimal placement of different RLXs considering the varying facility demands and pedestrian traffic.

The proposed optimisation model was designed to identify optimal RLX locations, ensuring pedestrian access to all areas neighbouring the railway line(s) while minimising the total weighted-walked (WD) distance. It is calculated by multiplying the distance by the number of pedestrians along a specific route. Several factors, such as environmental conditions and safety considerations, impact pedestrians' choices when they cross different rail sections using RLXs. Therefore, the objective function must consider the number of individuals to prioritise the routes with higher demand, as found in the demand-weighted vehicle routing problem for motorised transportation [29].

Consider an area of a mine with a network of m facilities, k locations as origins and l destinations (each one positioned at $b_k = (b_k^x, b_k^y)$). In addition, there are R rails dividing the area of interest into z zones with pedestrian traffic. These zones are determined by the rail adjacent to the facility downstream from the first to the penultimate one (see yellow arrow in Figure 1), while the last zone is given by $R + 1$. Given the mentioned conditions, this problem can be addressed as a continuous location problem, also known as the multi-facility Weber problem (MFWP), where the facilities can be freely located in a plane with an infinite set of potential locations [30]. Thus, all the RLXs can be set at any point on the rails. According to Hansen et al. [31], positioning the facilities within the convex hull of the customer set provides the optimal solution for the MFWP. Therefore, the facilities' locations coincide with the intersection points of the vertical and horizontal lines extended through the customer locations. Adapting the MFWP to the problem addressed in this work, the facilities are equivalent to the RLXs, and the customers are the facilities from which pedestrians circulate in the area of interest. Therefore, positioning at least one RLX on adjacent rails at the same X-coordinate of each facility results in the local optimum. Also, discretising the network based on the local optima of each facility reduces the problem's complexity. As a result, a network of $m \times R$ potential location points is obtained to place the RLXs. The model considers that pedestrians can move in all directions on a rectilinear path, so the network has no restrictions on the location of RLXs. Figure 1 shows an example of a grid with $m = 8$ facilities, $R = 4$ rails, and $z = 5$ zones.

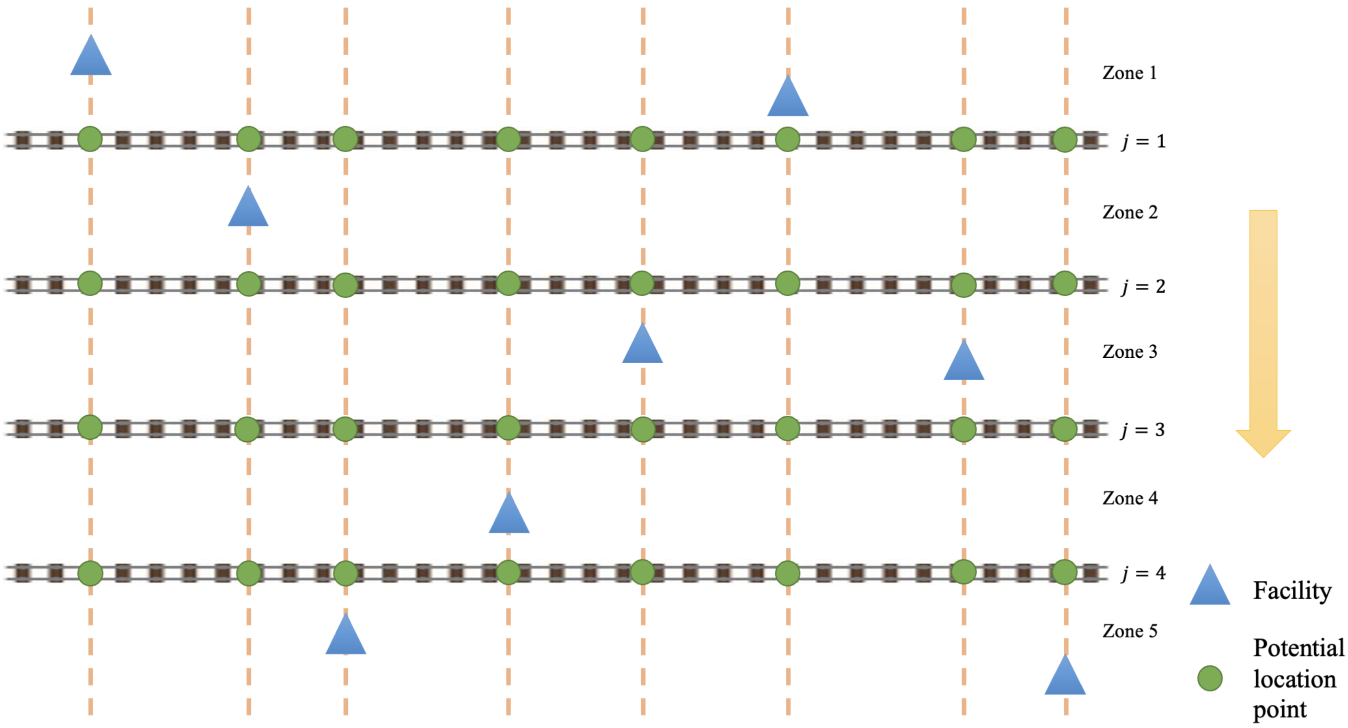


Figure 1: Example of a grid with $m = 8$ facilities, $R = 4$ rails, and $z = 5$ zones

It is suitable to mention that building an optimisation model to decide the location of level crossings is justified due to the combinatorial of possible k origins and l destinations (among the m facilities) for pedestrian circulation between the railways. In addition, the network of potential locations becomes higher as more facilities and railways are considered. Therefore, evaluating all the possible combinations will not provide an optimal solution within a reasonable time. The notation used for the model formulation is defined as follows:

Parameters

R : number of railways

n_j : set of potential points at railway j

m : number of facilities

$links$: set of feasible links between two facilities, where k is the origin and l is the destination and $k \neq l$

$a_{ij} = \left(a_{ij}^x, a_{ij}^y\right)$: coordinates of the potential point i at railway j

$b_k = (b_k^x, b_k^y)$: coordinates of the facility k

w_{kl} : number of pedestrians walking on the link (k, l)

p : number of RLXs to locate, where $p \geq R$

z_k : zone where facility k is located

Decision variables

X_{ij} : 1, if a RLX is located at potential point i of the railway j; 0, otherwise

Y_{ijkl} : 1, if the RLX located at potential point i of the railway j is assigned to the link (k, l); 0, otherwise

Thus, the mathematical model that determines the location of the RLXs to minimise the total WD can be expressed as:

$$
\begin{aligned}
Min\ Total\ WD \ = & \sum_{(k,l)\,\in\, links} \Bigg[\sum_{i=1}^{n_{j=z_k}} \left(\left|a_{i,j=z_k}^x - b_k^x\right| + \left|a_{i,j=z_k}^y - b_k^y\right|\right) . X_{i,j=z_k} . Y_{i,j=z_k,k,l} . w_{kl} \\
& + \sum_{i=1}^{n_j} \sum_{u=1}^{n_{j+1}} \sum_{j=z_k}^{z_l-2} \left(\left|a_{i,j}^x - a_{u,j+1}^x\right| + \left|a_{i,j}^y - a_{u,j+1}^y\right|\right) . X_{ij} . X_{u,j+1} . Y_{ijkl} . Y_{u,j+1,k,l} . w_{kl} \\
& + \sum_{i=1}^{n_{j=z_l-1}} \left(\left|b_l^x - a_{i,j=z_l-1}^x\right| + \left|b_l^y - a_{i,j=z_l-1}^y\right|\right) . X_{i,j=z_l-1} . Y_{i,j=z_l-1,k,l} . w_{kl} \Bigg]
\end{aligned} \tag{1}
$$

Subject to

$$
\sum_{j=1}^{R} \sum_{i=1}^{n_j} X_{ij} = p \tag{2}
$$

$$\sum_{i=1}^{n_j} X_{ij} \geq 1 \qquad \forall\, j \in R \tag{3}$$

$$Y_{i,j=z_k,k,l} \leq X_{i,j=z_k} \qquad \forall\, i \in n_{j=z_k}\,;\, (k,l) \in links \tag{4}$$

$$Y_{ijkl} \leq X_{ij} \qquad \forall\, i \in n_j\,;\, j = z_k, \ldots, z_l - 1\,;\, (k,l) \in links \tag{5}$$

$$Y_{i,j=z_l-1,k,l} \leq X_{i,j=z_l-1} \qquad \forall\, i \in n_{j=z_l-1}\,;\, (k,l) \in links \tag{6}$$

$$\sum_{i=1}^{n_{j=z_k}} Y_{i,j=z_k,k,l} = 1 \qquad \forall\, (k,l) \in links \tag{7}$$

$$\sum_{i=1}^{n_j} Y_{ijkl} = 1 \qquad \forall\, j = z_k, \ldots, z_l - 1\,;\, (k,l) \in links \tag{8}$$

$$\sum_{i=1}^{n_{j=z_l-1}} Y_{i,j=z_l-1,k,l} = 1 \qquad \forall\, (k,l) \in links \tag{9}$$

$$X_{ij} \in \{0,1\}\,;\, Y_{ijkl} \in \{0,1\} \qquad \forall\, i,u \in n_j\,;\, j \in R\,;\, (k,l) \in links \tag{10}$$

The objective function (1) minimises the total WD considering the number of pedestrians, all the links (k,l) in the network and the paths (a) from the origin to the first railway, (b) between railways and (c) the last railway to the destination. Constraint (2) states that the total number of RLXs to be located must be p while constraints (3) assure that each railway has at least one RLX. Constraints (4), (5), and (6) ensure that a RLX is allocated as long as it has been placed on all the railways to be crossed for the link (k,l). Constraints (7), (8), and (9) determine that only one RLX per railway can be allocated for the link (k,l). Finally, constraints (10) specify the domain of the decision variables.

3. Proposed genetic algorithm

The genetic algorithm (GA) is a widely used search technique for solving combinatorial optimisation problems [32]. It mimics Charles Darwin's theory of natural selection and survival of the fittest individuals. GAs work with encoded solutions named *population*. Each solution is represented by a string of genes called *chromosome*. In addition, each one has an objective function value called *fitness*. The current population is modified using genetic operators: *crossover* and *mutation*, resulting in a set of *offspring*. The evolution procedure is performed generation by generation until a stop criterion is reached. The overall structure of the GA is described as follows:

1. *Parameter setting:* set the population size (pop_size), crossover rate ($prob\_cross$), mutation rate ($prob\_mut$), and number of generations (gen_total).
2. *Initial population:* create an initial population of pop_size chromosomes using random generation and a proposed algorithm to correct feasibility.
3. *Offspring generation*: compute the fitness for each chromosome and recombine the individuals of the current population to create the offspring using the crossover and mutation operators. New individuals are generated until a fixed number of chromosomes are reached.
4. *Current best solution*: compute the fitness value (total WD) for each offspring and select the best solution from the enlarged population (current + offspring).
5. *Replacement*: select pop_size solutions from the enlarged population to determine the next generation.
6. *Stop criterion*: the GA stops if a predefined number of generations have been executed. When the stop criterion is satisfied, the GA returns the best solution as output; otherwise, it repeats steps 3-5.

The following subsections describe in detail the different steps of the GA.

3.1. Chromosome codification and initial population

The chromosome is represented by a binary-coded matrix of size $R \times \max(n_j)$ that describes the location of the p RLXs. In this array, the number 1 represents that an RLX is located at the potential point $i \in n_j$ at the railway $j \in R$, while the number 0 shows the potential points where an

RLX is not placed. Figure 2 presents a chromosome for a network with five railways ($R = 5$) and 17 RLXs ($p = 17$).

Rail 1	0	1	0	0	0	0	1
Rail 2	0	1	0	1	1	0	1
Rail 3	0	0	1	1	0	0	1
Rail 4	0	1	1	1	1	0	1
Rail 5	0	0	1	0	1	0	1

Figure 2: Example of a grid with $R = 5$ railways, $p = 17$ RLXs, and $n_1 = n_2 = n_3 = n_4 = n_5 = 7$ potential points per rail

The GA starts by generating an initial population of chromosomes as a basis for finding a good solution by applying the crossover and mutation operators. The initial population's chromosomes were randomly created as a diversification approach that avoids falling into local optima. However, these chromosomes may represent infeasible solutions, i.e., the total number of RLXs differs from p, and a rail does not have RLXs located at potential points. Therefore, a heuristic was used to correct the chromosomes with infeasible solutions, whose procedure is shown in Algorithm 1.

It starts by calculating the number of RLXs located in each rail (lines 1-5). If any rail is empty, a random number of RLXs to be located at random positions on that rail is determined. Then, the total number of RLXs is verified to be the value of p (lines 6-20). If $\sum_{j=1}^{R} \sum_{i=1}^{n_j} X_{ij} > p$, the necessary RLXs are eliminated on rails having more than one RLX. Conversely, if $\sum_{j=1}^{R} \sum_{i=1}^{n_j} X_{ij} < p$, the necessary RLXs are added on rails having less than n_j RLXs. The result of this heuristic is a feasible chromosome (line 21). Thus, the fitness value is calculated for each link (k, l) by assigning the RLXs that result in the lowest total WD.

3.2. Population improvement and termination criteria

As was mentioned, the GA is a bio-inspired metaheuristic based on the natural selection theory in which the crossover and mutation operator generate new individuals. The matrix-structured chromosome is turned into a vector to facilitate the application of the genetic operators. For the crossover, two chromosomes (parent #1 and parent #2) are randomly selected from the current population. Two genes within parent #1 are randomly selected, and the sub-string between them is copied to the offspring in the corresponding positions. The remaining offspring positions are

Algorithm 1 Feasibility check

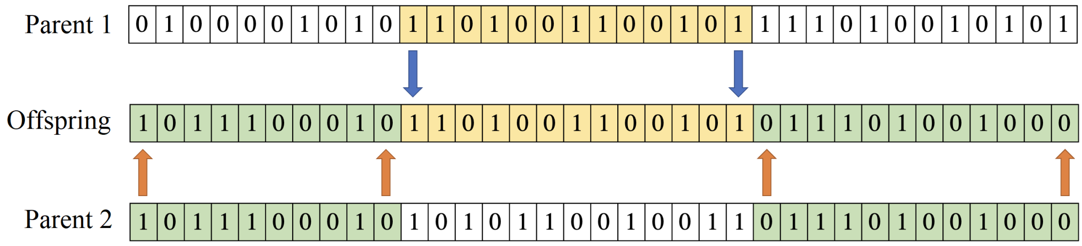
for $j \in R$ **do**
    if $\sum_{i=1}^{n_j} X_{ij} = 0$ **then**
        Add a random number of RLXs (between 1 and n_j) and place them randomly on rail j
    end if
end for
Calculate $T = \sum_{j=1}^{R} \sum_{i=1}^{n_j} X_{ij} - p$
if $T > 0$ **then**
    repeat
        Let $R^* \in R$ the set of rails where $\sum_{i=1}^{n_j} X_{ij} > 1$
        Randomly select a rail j^* from R^* and a potential point i^* with $X_{i^*,j^*} = 1$ from j^*
        $X_{i^*,j^*} \leftarrow 0$
        $T = T - 1$
    until $T = 0$
end if
if $T < 0$ **then**
    repeat
        Let $R^* \in R$ the set of rails where $\sum_{i=1}^{n_j} X_{ij} < n_j$
        Randomly select a rail j^* from R^* and a potential point i^* with $X_{i^*,j^*} = 0$ from j^*
        $X_{i^*,j^*} \leftarrow 1$
        $T = T + 1$
    until $T = 0$
end if
return the feasible chromosome

copied from parent #2, making a left-to-right scan. Figure 3a shows an example of the crossover operator. For the mutation operator, a chromosome (parent) is randomly selected from the current population. Two positions (with different binary codes) within the parent are randomly selected. Then, the offspring is built by inverting the genes. Figure 3b presents an example of the mutation operator.

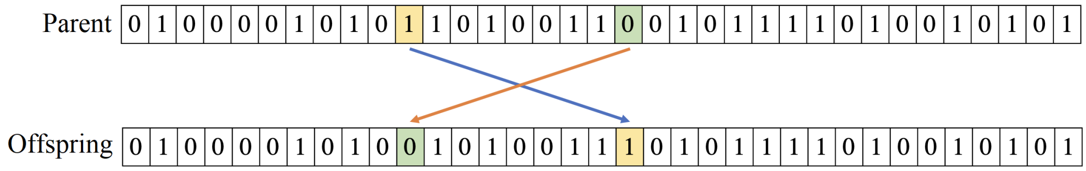
If the crossover and mutation operators yield infeasible solutions, the heuristic described in Algorithm 1 must be applied to correct those chromosomes. The next population is built by selecting individuals from the current population and the new offspring until *pop_size* chromosomes are reached. The overall process is repeated until the total number of generations (*gen_total*) is executed.

4. Computational experiments and results

This section describes the test instances, the experimental design to select the best parameters of the GA, and a sensitivity analysis to study the impact of the railway network on the total WD.



(a) An example of the crossover operator



(b) An example of the mutation operator

Figure 3: Example of the genetic operators

4.1. Data generation

The effectiveness of the GA must be tested under different network conditions (number of railways, facilities, and number of pedestrians) as they determine the complexity and size of the problem. However, no test instances were found in the literature since the literature does not have studies focused on the strategic placement of RLXs in a railway network. Therefore, ten (10) instances were generated, as shown in Table 1.

Table 1: Test instances

Problem	R	m	w_{kl}	p
P01	2	4	U[35,50]	4
P02	4	5	U[35,60]	10
P03	4	8	U[20,60]	16
P04	4	10	U[5,35]	20
P05	5	10	U[5,35]	25
P06	5	7	U[5,30]	17
P07	5	15	U[1,30]	35
P08	6	9	U[5,30]	27
P09	6	18	U[1,30]	51
P10	7	25	U[1,30]	66

All instance data were determined based on the characteristics of the area under study described in Section 2. The current system comprises four rail lines; however, this number may vary depending on the mining operation's requirements. The number of facilities was set according to the quantity of administrative and maintenance zones within the rail lines. The number of pedestrians was calculated based on the daily personnel arrival rate. It considers individuals departing from the different administrative and maintenance areas who cross the railroad lines to reach their destinations. This data collection was conducted throughout one week of operational days. The information for the rest of the parameters can be found at `https://github.com/jubizm/Level_crossings_location.git`.

4.2. Parameter tuning

The literature has shown that metaheuristics' performance highly depends on the choice of their parameters. Therefore, a 2^4 factorial design was used to analyse the influence of the key parameters on the objective function. Each factor was studied with two levels, and the experiment was repeated 9 times for a total of 160 runs per instance. The selected evaluation indicator is the objective function of the model, i.e., the total WD. The considered factors and their levels are: (1) *pop_size* tested at {50, 300}, (2) *prob_cross* tested at {0.5, 0.8}, (3) *prob_mut* tested at {0.1, 0.4}, and (4) *gen_total* tested at {100, 500}. Table 2 shows the level of each factor per run.

Table 2: 2^4 factorial design

Experiment number	*pop_size*	*prob_cross*	*prob_mut*	*gen_total*
1	300	0.8	0.4	500
2	50	0.5	0.4	500
3	50	0.5	0.4	100
4	300	0.5	0.1	100
5	50	0.5	0.1	500
6	50	0.5	0.1	100
7	50	0.8	0.1	100
8	50	0.8	0.4	100
9	300	0.5	0.1	500
10	300	0.8	0.4	100
11	300	0.5	0.4	100
12	300	0.8	0.1	100
13	50	0.8	0.4	500
14	300	0.8	0.1	500
15	300	0.5	0.4	500
16	50	0.8	0.1	500

Computational experiments were performed on an Apple M1 Pro with 16 GB of RAM, and the GA was coded and compiled in Matlab R2022a. The significant factors were identified by comparing the P-value given by the analysis of variance (ANOVA) with a significance level $\alpha = 0.05$. Those factors with a P-value less than 0.05 are relevant, i.e., they highly influence the minimisation of the total WD. Table 3 summarises the results for problems P03 to P10, where the significant factors are denoted with a check mark. The results of all runs for problems P01 and P02 were the same, i.e., all factors can be set at any level to minimise the objective function.

Table 3: Significant factors for each problem

Problem	*pop_size*	*prob_cross*	*prob_mut*	*gen_total*
P03	✓		✓	✓
P04	✓	✓		✓
P05	✓			✓
P06	✓	✓		✓
P07	✓		✓	✓
P08	✓			✓
P09	✓		✓	✓
P10	✓		✓	✓

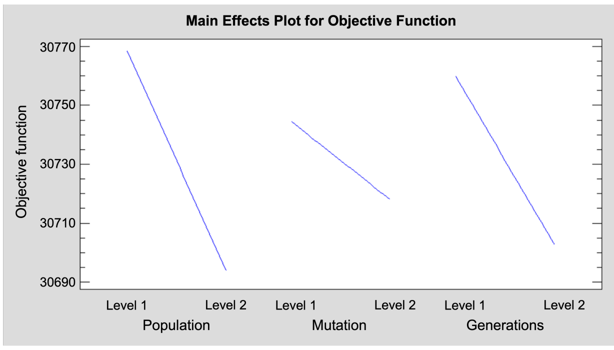
Results show that *pop_size* and the *gen_total* are the most relevant factors in all problems (small and large). Moreover, the main effects plots indicate that these two factors highly impact minimising the objective function, as shown in Figure 4.

This means that the higher the value set for each, the lower the total weighted distance. On the other hand, the crossover rate was significant only in problems P04 and P06, while the mutation rate was in P03, P07, P09 and P10. These findings are statistically valid since normality, constant variance, and independence assumptions were validated with satisfactory results. The statistical tests performed were Shapiro Wilk, Barlett and Durbin Watson [33]. Table 4 shows the results of the P-values for each test. They must be greater than 0.05 to meet each assumption.

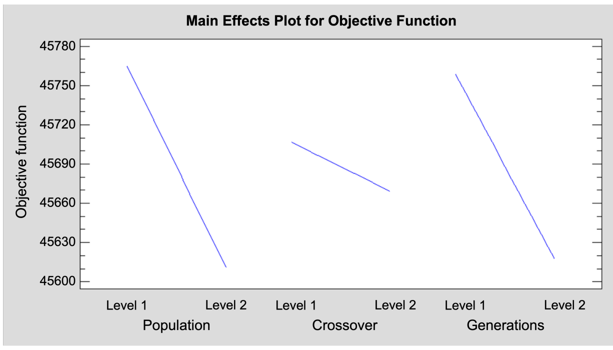
Table 4: Assumption validation results

Problem	Normality test	Variance check				Independence test
		pop_size	*prob_cross*	*prob_mut*	*gen_total*	
P03	0.2474	0.8934	0.1793	0.4190	0.6219	0.1234
P04	0.3657	0.1762	0.9747	0.3178	0.1071	0.6492
P05	0.1237	0.1241	0.6783	0.9644	0.1101	0.2480
P06	0.0743	0.0616	0.5416	0.1991	0.3600	0.4912
P07	0.4431	0.2879	0.2397	0.6925	0.3199	0.6037
P08	0.1836	0.0834	0.1780	0.0874	0.4224	0.1612
P09	0.9334	0.0814	0.8641	0.8656	0.6594	0.3843
P10	0.9382	0.0644	0.4111	0.3880	0.5153	0.2156

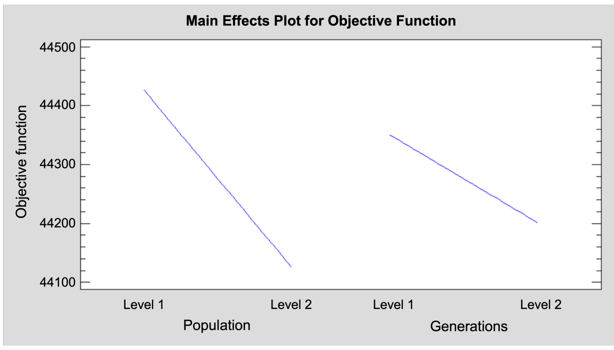
Based on the results of the experiments, the optimal factor levels were calculated using the method of steepest descent. The best parameter setting of the GA per problem and the 95% confidence



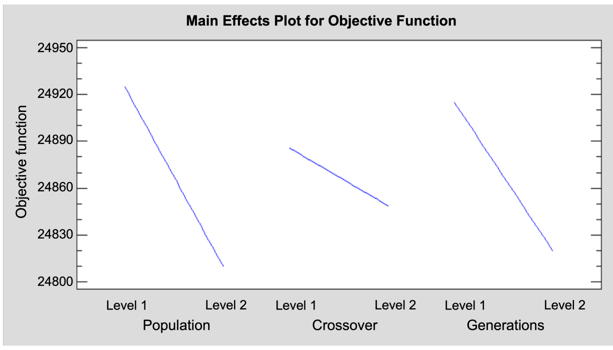
(a) Main effects plot for problem P03



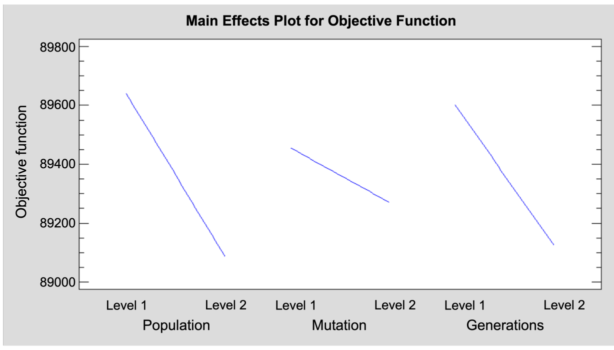
(b) Main effects plot for problem P04



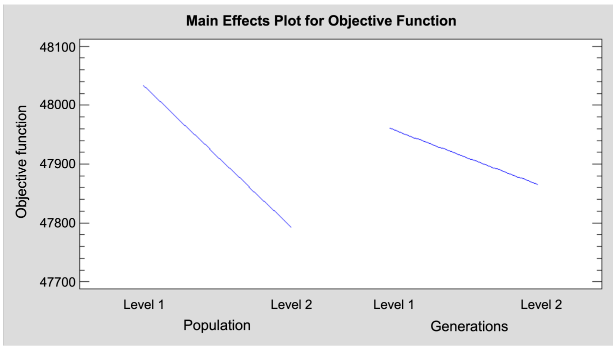
(c) Main effects plot for problem P05



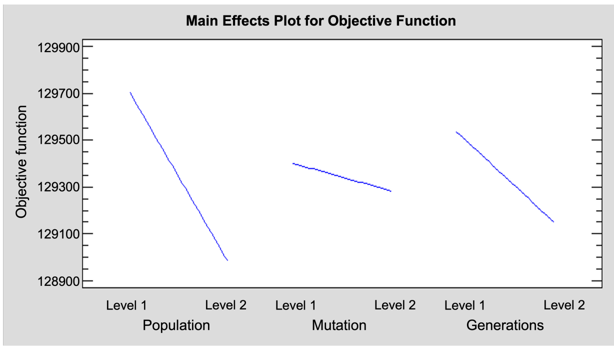
(d) Main effects plot for problem P06



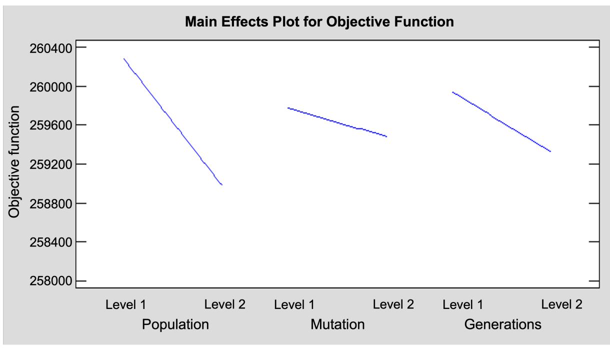
(e) Main effects plot for problem P07



(f) Main effects plot for problem P08



(g) Main effects plot for problem P09



(h) Main effects plot for problem P10

Figure 4: Main effects for the total WD per problem

interval for the objective function are shown in Table 5.

Table 5: Optimum levels of factor levels and 95% confidence interval for the objective function

Problem	*pop_size*	*prob_cross*	*prob_mut*	*gen_total*	Lower confidence limit	Mean WD	Upper confidence limit
P03	297	0.7594	0.4	500	30,624.9	30,642	30,659.6
P04	295	0.5	0.1	500	45,460.8	45,508	45,554.2
P05	300	0.508	0.2336	500	44,011.5	44,051	44,089.7
P06	300	0.8	0.2466	500	24,700.2	24,726	24,750.8
P07	300	0.7256	0.4	500	88,644.6	88,756	88,866.7
P08	300	0.7187	0.1	500	47,714.2	47,737	47,759.6
P09	300	0.7579	0.4	459	128,634	128,739	128,843
P10	300	0.5001	0.3609	500	258,358	258,554	258,751

Using the calibrated parameters, the location (X_{ij}) of RLXs and their assignment (Y_{ijkl}) to each link (k, l) per problem were determined.

4.3. Results

As mentioned above, the location and assignment of RLXs for all the problems were calculated using the calibrated GA parameters. Table 6 presents the number of links, number of RLXs to located, assigned RLXs, free RLXs and total WD.

Table 6: Results per problem

Problem	Number of *links*	p	Number of assigned RLXs	Number of free RLXs	WD
P01	5	4	2	2	3,826
P02	10	10	6	4	10,898
P03	26	16	15	1	30,654
P04	38	20	18	2	45,533
P05	40	25	20	5	44,048
P06	19	17	15	2	24,711
P07	92	35	29	6	88,644
P08	34	27	18	9	47,751
P09	136	51	38	13	128,729
P10	269	66	49	17	258,449

The results indicate that not all the located RLXs are assigned to links because one RLX can connect to more than one origin-destination pair. Therefore, the total WD is reduced, prioritising those links with higher pedestrians. However, the unassigned RLXs remain as backups. Thus, the optimal solution can be modified in case of any system change without including new RLXs. Two examples can be seen in Figures 5 and 6 for problems P02 and P07, respectively.

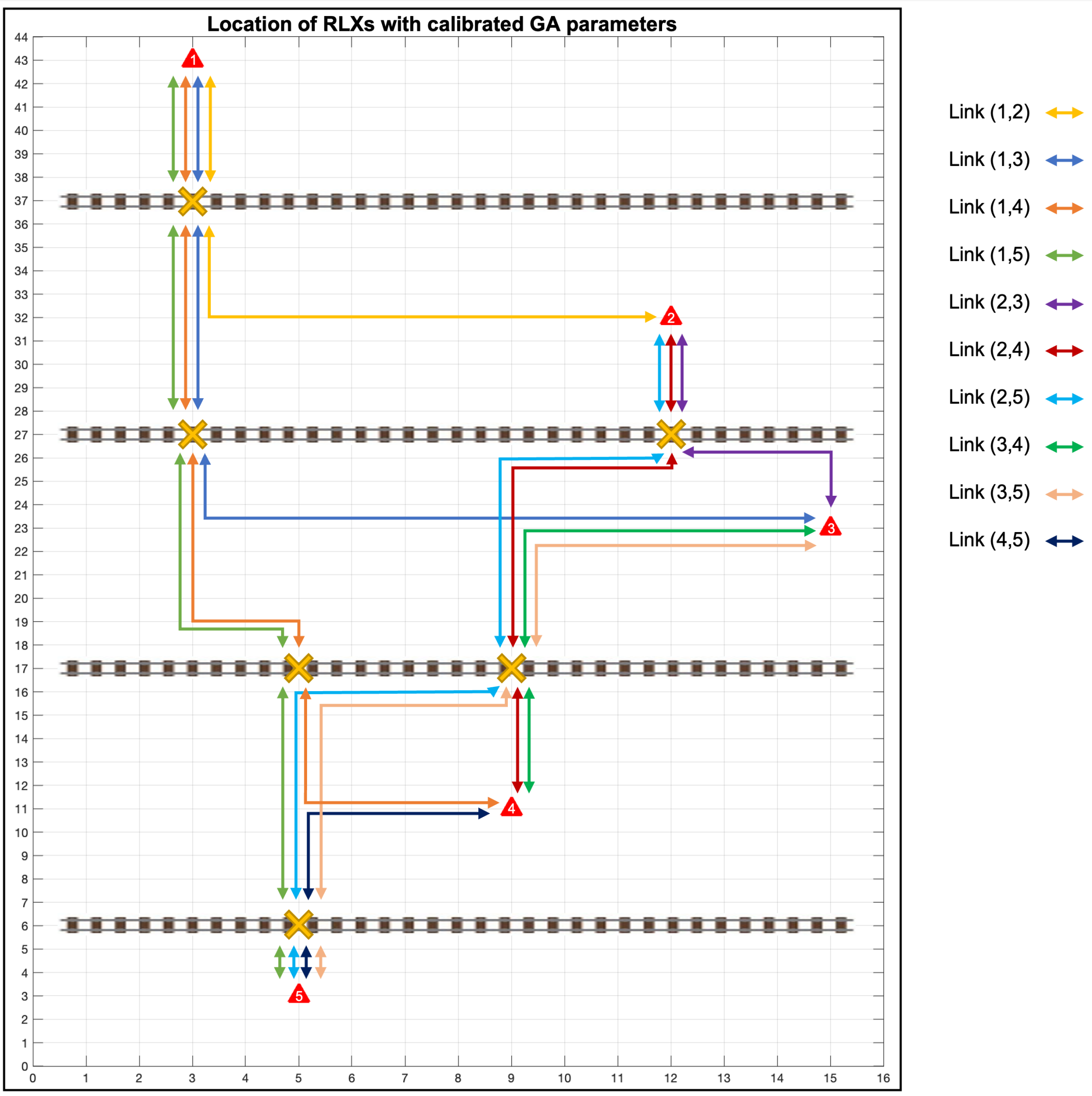


Figure 5: Location of RLXs for problem P02

Figure 5 shows the bidirectional circulation routes connecting the five facilities to obtain a total WD of 10,898. The colours show the connection between rails and RLXs to be travelled by the pedestrians of each link (k, l). The results indicate that, for problem P02, an RLX must be installed on rail $j = 1$ at the exit of facility $k = 1$. Two RLXs should be located on rail $j = 2$: one to connect it with rail $j = 1$ and the other at the exit of facility $k = 2$. Two RLXs should be installed on rail $j = 3$ to connect the routes from rail $j = 2$ (including the ones from $k = 3$) and to the arrival of facility $k = 4$. Finally, an RLX should be placed on rail $j = 4$ to connect with facility $k = 5$.

For the other problems, the solution presents coloured sections indicating all the segments pedestrians must travel to complete the route per link (k, l). The black dotted lines correspond to segments common to several routes of different links. Figure 6 shows that, for problem P07, 4 RLXs should be

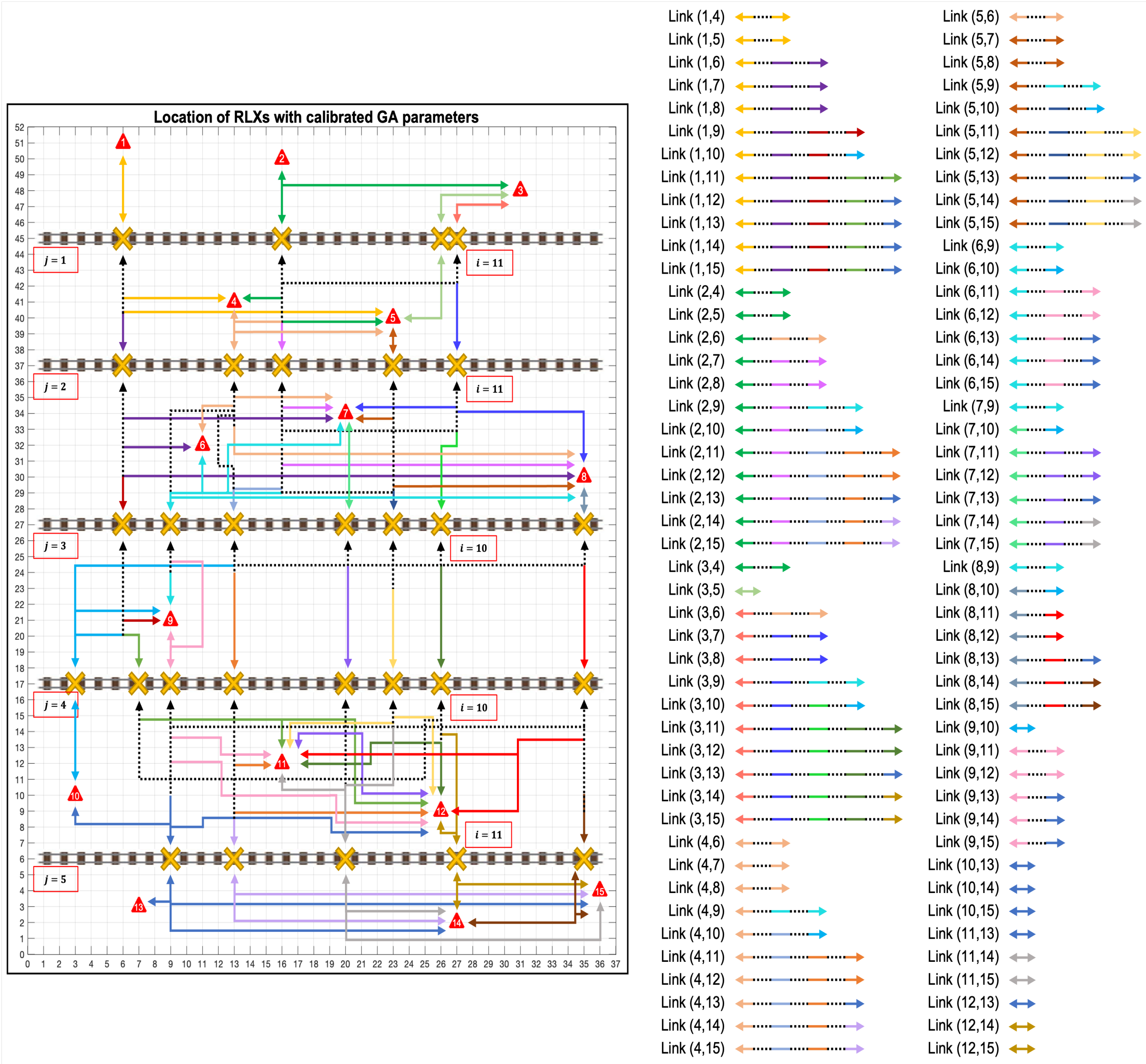


Figure 6: Location of RLXs for problem P07

located on rail $j = 1$, 5 RLXs on rail $j = 2$, 7 RLXs on rail $j = 3$, 8 RLxs on rail $j = 4$ and 5 RLXs on rail $j = 5$. The total weighted distance achieved with the solution is 88,644. An example is the path of the link $(3, 15)$. The route that a pedestrian should travel is facility $k = 3 \to$ pink segment $\to$ RLX at $i = 11, j = 1 \to$ **black segment** $\to$ blue segment $\to$ RLX at $i = 11, j = 2 \to$ **black segment** $\to$ light green segment $\to$ RLX at $i = 10, j = 3 \to$ **black segment** $\to$ dark green segment $\to$ RLX at $i = 10, j = 4 \to$ **black segment** $\to$ brown segment $\to$ RLX at $i = 11, j = 5$ $\to$ facility $k = 15$. The same structure is used to present the rest of the routes. It is important to highlight that several of the located RLXs are aligned with the X-coordinate of the facilities, either at the entrance or exit, as mentioned in Section 2. However, it can be observed how the size of the problem grows exponentially, i.e., the number of links for which the circulation route must be determined. It justifies the need to add more RLXs trying to cover several links without increasing the total WD for pedestrians.

4.4. Sensitivity analysis

As mentioned above, not all the located RLXs were assigned to routes since a single RLX can serve several links. However, these free RLXs can be used in case of system changes. This section analyses the impact of each facility's position on the location and assignment of RLXs in the rail network. Using the calibrated GA parameters mentioned in Table 5, the assignment for each feasible b_k^x per facility to both the left and right of the initial location was determined. This procedure was conducted by moving one facility at a time, i.e., the other facilities remain at their original b_k^x. Based on the results, the range of b_k^x in which the location and assignment of RLXs remain the same could be established. It is useful to decide where to locate new facilities without changing the optimal solution associated with the location problem. Results for problems P02 and P08 are presented in Figures 7 and 8, where 10 and 27 RLXs were located, respectively. The green lines show the b_k^x where the assignment does not change, the orange lines indicate the b_k^x that modifies the assignment, and the grey lines determine the unfeasible b_k^x.

Some facilities are more restricted than others. For example, facility 4 in Figure 7 can move in both directions without changing the RLX assignment, while facility 1 can move only two positions to the left. The findings in Figure 8 also have this feature. The range of b_k^x values depends on the initial position and the number of RLXs to be located. If a larger number of RLXs are placed, the original solution is more susceptible to change because the possible combinations of assigned RLXs increase. Moreover, each RLX is used in fewer links. However, as the system grows, the smaller the variation range of b_k^x, where the RLX assignment remains the same. It means that the proposed methodology is flexible to changes in the network conditions, i.e., it is sensitive to the problem's size. Results of the analysis for the rest of the problems are described in Appendix A and Appendix B.

5. Conclusions

RLX accidents have significant safety and financial concerns due to human and material losses. According to statistics from the United States and Europe, collisions at RLXs are a leading cause of railway fatalities globally. This problem can be found in urban and industrial environments, such as

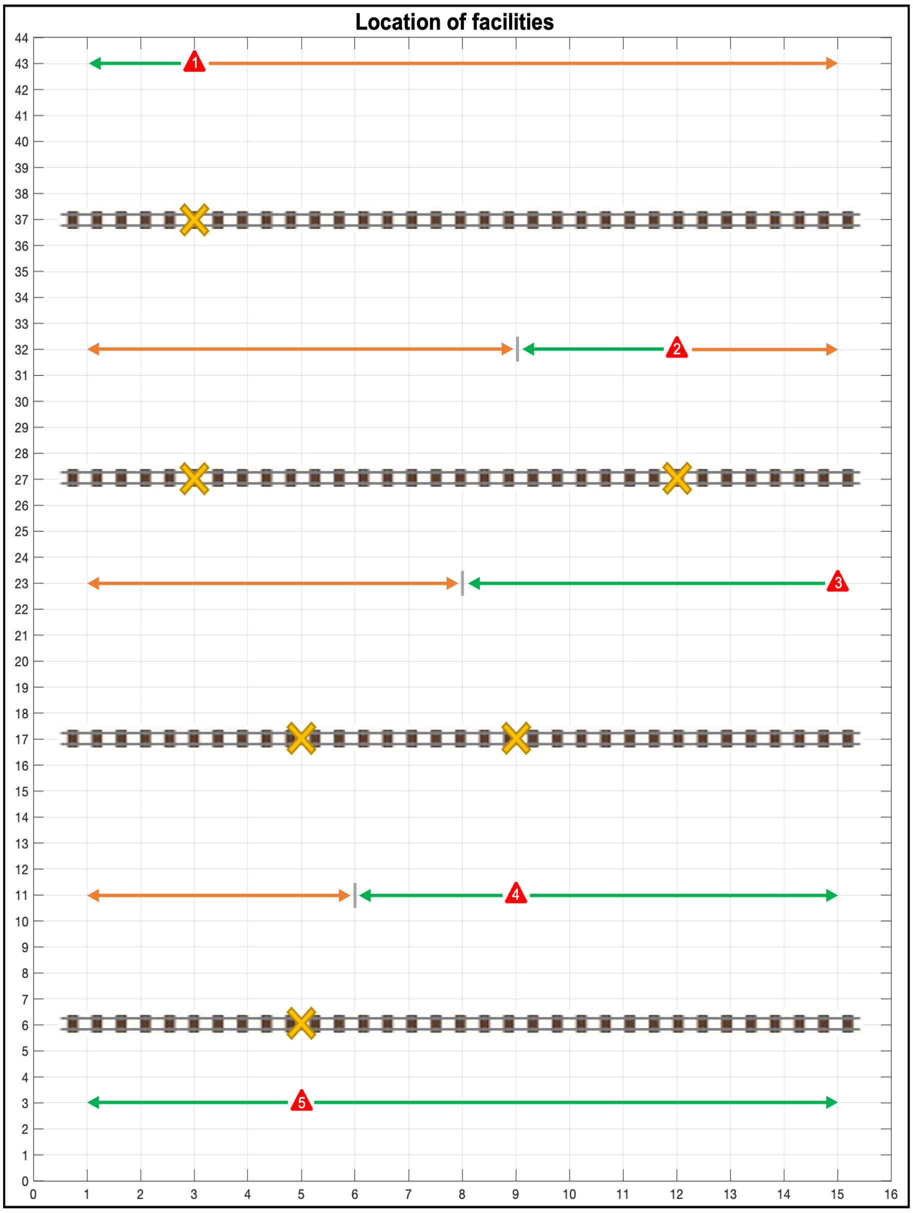


Figure 7: Variation range of b_k^x per facility for problems P02

mining, where rail transport is pivotal because of its advantages in terms of cost and efficiency. However, the system's characteristics still challenge stakeholders' decision-making process. On the other hand, unsafe pedestrian behaviours contribute significantly to RLX accidents, impacted by convenience and different risk factors. Literature highlights the importance of RLX location in influencing dangerous behaviours, accident severity, and violation rates, suggesting the need for a systematic approach to RLX placement. Nevertheless, existing research lacks quantitative methodologies based on optimisation techniques to determine RLX locations in railway systems. Therefore, this study develops an mathematical model to locate RLXs on a railway system to minimise the total weighted-

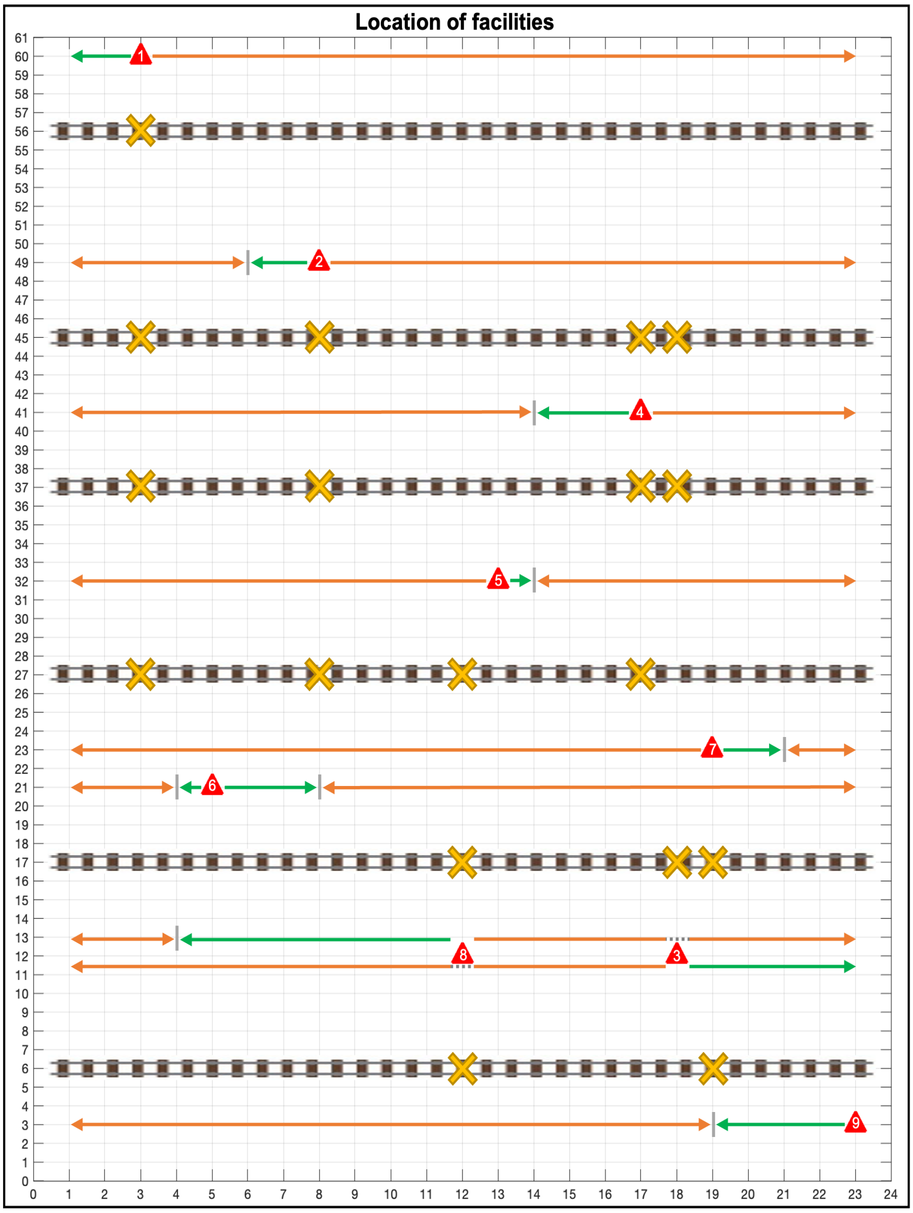


Figure 8: Variation range of b_k^x per facility for problems P08

walked distance. It is an adapted MFWP, where the facilities are the RLXs, and the customers are the facilities from which the pedestrians travel in the zone under study. In addition, a GA was developed to solve the mathematical model for any problem size. Its parameters and performance were evaluated using ten instances generated based on the system's characteristics.

The results show that the model seeks to locate the RLXs at the entrance or exit of the facilities, prioritising the routes with the highest pedestrian demand. In addition, not all the RLXs on the rail lines are assigned to the circulation routes. The number of links ranged from 5 to 269, and the percentage of RLXs allocated was between 50% and 94%. It means that the unassigned RLXs are left to support the system in case of disruptive events that force the use of alternate routes without significantly increasing the total weighted-walked distance. For this reason, a sensitivity analysis was performed to show the model's flexibility in response to system changes. The facilities' locations were progressively modified to determine how far the optimal allocation and routing changed. The results showed that, as the problem size grows and more RLXs are placed, the optimal solution is more susceptible to change. In addition, each RLX is assigned to fewer routes because pedestrians have more crossing points.

In general, this study provides a methodology to determine the location of RLXs in rail systems to optimise metrics that play a fundamental role in pedestrian decisions. As was mentioned, they tend to choose the shortest path because it involves less time and effort. Therefore, locating RLXs is a decision that must be supported by criteria related to pedestrian safety. It is suitable to mention that this type of proposal is necessary because an RLX can be at any point on a rail, in addition to all the possible combinations of routes that a pedestrian can use. This is why it has become a tool that supports infrastructure decisions for stakeholders in the mining industry but can also be transferred to other contexts, such as urban environments.

Finally, future research should include a cost-benefit analysis related to the types of RLX, collision costs (insurance payments), maintenance costs, and budget constraints. Moreover, it is suitable to consider other indicators to measure the risk while crossing and the climate conditions that affect pedestrians' decisions. In addition, the proposed methodology can be used in other environments such as urban systems where pedestrians must cross roads.

References

[1] L.-S. Tey, L. Ferreira, A. Wallace, Measuring driver responses at railway level crossings, Accident Analysis & Prevention 43 (2011) 2134–2141.

[2] B. Lobb, Trespassing on the tracks: A review of railway pedestrian safety research, Journal of Safety Research 37 (2006) 359–365.

[3] Federal Railroad Administration, Highway-rail grade crossings overview, 2021.

[4] Eurostat Statistic Explained, Railway safety statistics in the eu, 2022.

[5] C. Lo, On track: five essential mining rail links in development, 2013.

[6] Deloitte Access Economics, Value of rail 2020. the rail industry's contribution to a strong economy and vibrant communities, 2020.

[7] Ministerio de Transporte, Portal logístico de colombia, 2020.

[8] Mining Safety, Rail/level crossings and mining safety, 2021.

[9] B. Lobb, N. Harre, T. Suddendorf, An evaluation of a suburban railway pedestrian crossing safety programme, Accident Analysis & Prevention 33 (2001) 157–165.

[10] I. Savage, Trespassing on the railroad, Research in Transportation Economics 20 (2007) 199–224.

[11] J. Freeman, A. Rakotonirainy, Mistakes or deliberate violations? a study into the origins of rule breaking at pedestrian train crossings, Accident Analysis & Prevention 77 (2015) 45–50.

[12] T. Tjahjono, A. Kusuma, Y. Y. Pratiwi, R. Y. Purnomo, Identification determinant variables of the injury severity crashes at road-railway level crossing in indonesia, Transportation research procedia 37 (2019) 211–218.

[13] D. Barić, G. M. Havârneanu, C. Măirean, Attitudes of learner drivers toward safety at level crossings: do they change after a 360° video-based educational intervention?, Transportation research part F: traffic psychology and behaviour 69 (2020) 335–348.

[14] T. Stefanova, O. Oviedo-Trespalacios, J. Freeman, C. Wullems, A. Rakotonirainy, J.-M. Burkhardt, P. Delhomme, Contextual factors explaining risk-taking intentions at australian level crossings, Safety science 110 (2018) 145–161.

[15] P. Blaho, L. Pečenỳ, J. Gašparík, Causality of accidents at railway-crossings in slovakia and its prevention, in: 2020 XII International Science-Technical Conference AUTOMOTIVE SAFETY, IEEE, pp. 1–6.

[16] W. Hao, C. Kamga, D. Wan, The effect of time of day on driver's injury severity at highway-rail grade crossings in the united states, Journal of traffic and transportation engineering (English edition) 3 (2016) 37–50.

[17] Z. Zheng, P. Lu, D. Pan, Predicting highway–rail grade crossing collision risk by neural network systems, Journal of Transportation Engineering, Part A: Systems 145 (2019) 04019033.

[18] L.-S. Tey, L. Ferreira, Driver compliance at railway level crossings, in: 33rd Australasian Transport Research Forum.

[19] M. G. Lenné, C. M. Rudin-Brown, J. Navarro, J. Edquist, M. Trotter, N. Tomasevic, Driver behaviour at rail level crossings: Responses to flashing lights, traffic signals and stop signs in simulated rural driving, Applied ergonomics 42 (2011) 548–554.

[20] G. J. Read, P. M. Salmon, M. G. Lenné, N. A. Stanton, Walking the line: Understanding pedestrian behaviour and risk at rail level crossings with cognitive work analysis, Applied ergonomics 53 (2016) 209–227.

[21] P. M. Salmon, G. J. Read, N. A. Stanton, M. G. Lenné, The crash at kerang: Investigating systemic and psychological factors leading to unintentional non-compliance at rail level crossings, Accident Analysis & Prevention 50 (2013) 1278–1288.

[22] V. Beanland, P. M. Salmon, A. J. Filtness, M. G. Lenné, N. A. Stanton, To stop or not to stop: Contrasting compliant and non-compliant driver behaviour at rural rail level crossings, Accident Analysis & Prevention 108 (2017) 209–219.

[23] D. Barić, H. Pilko, M. Starčević, Introducing experiment in pedestrian behaviour and risk perception study at urban level crossing, International journal of injury control and safety promotion 25 (2018) 102–112.

[24] G. J. Read, J. A. Cox, A. Hulme, A. Naweed, P. M. Salmon, What factors influence risk at rail level crossings? a systematic review and synthesis of findings using systems thinking, Safety science 138 (2021) 105207.

[25] S. Anandarao, C. D. Martland, Level crossing safety on east japan railway company: Application of probabilistic risk assessment techniques, Transportation 25 (1998) 265–286.

[26] J. Davey, A. Wallace, N. Stenson, J. Freeman, The experiences and perceptions of heavy vehicle

drivers and train drivers of dangers at railway level crossings, Accident Analysis & Prevention 40 (2008) 1217–1222.

[27] B. Palat, F. Paran, P. Delhomme, Applying an extended theory of planned behavior to predicting violations at automated railroad crossings, Accident Analysis & Prevention 98 (2017) 174–184.

[28] C. Liang, M. Ghazel, O. Cazier, E.-M. El-Koursi, Analyzing risky behavior of motorists during the closure cycle of railway level crossings, Safety science 110 (2018) 115–126.

[29] J. D. Camm, M. J. Magazine, S. Kuppusamy, K. Martin, The demand weighted vehicle routing problem, European Journal of Operational Research 262 (2017) 151–162.

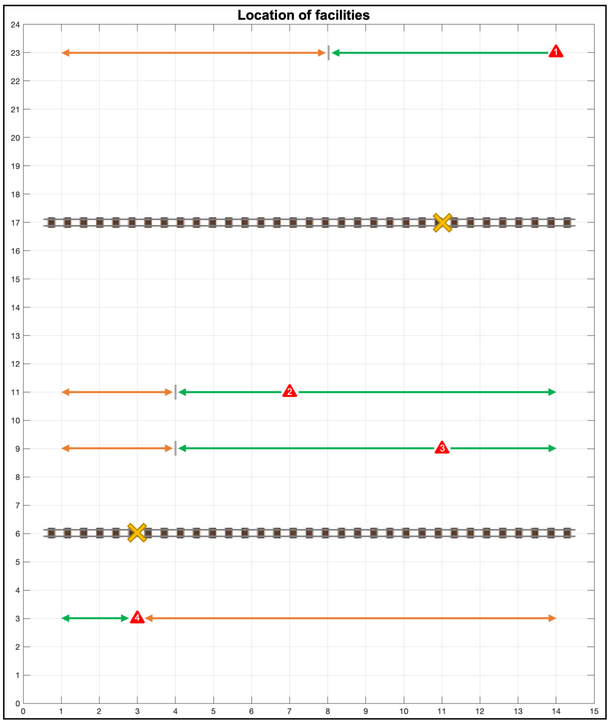
[30] S. T. W. Mara, R. Kuo, A. M. S. Asih, Location-routing problem: a classification of recent research, International Transactions in Operational Research 28 (2021) 2941–2983.

[31] P. Hansen, J. Perreur, J.-F. Thisse, Location theory, dominance, and convexity: Some further results, Operations Research 28 (1980) 1241–1250.

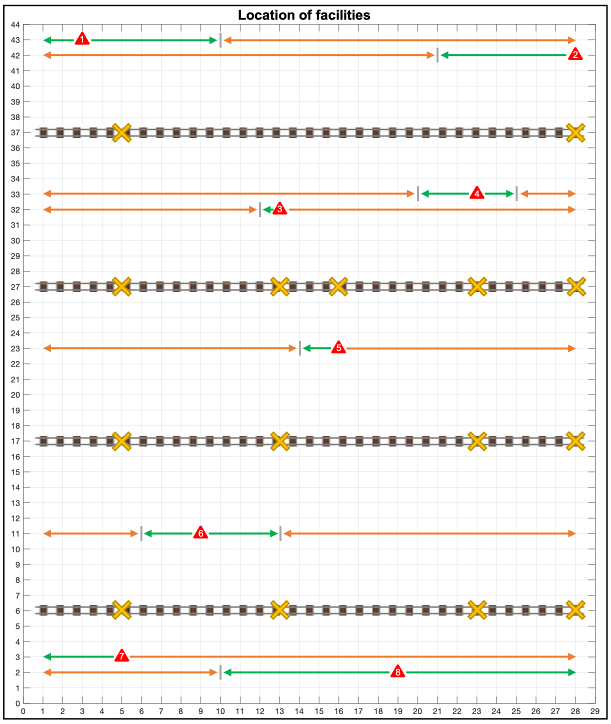
[32] J. H. Holland, et al., Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence, MIT press, 1992.

[33] D. Montogomery, Design and Analysis of Experiments, Wiley, 2012.

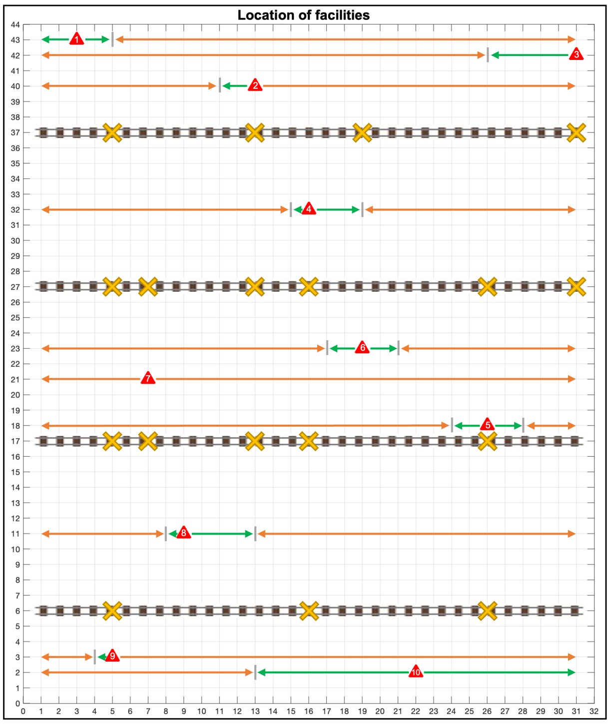
Appendix A. Results of sensitivity analysis for problems P01, P03, P04 and P05



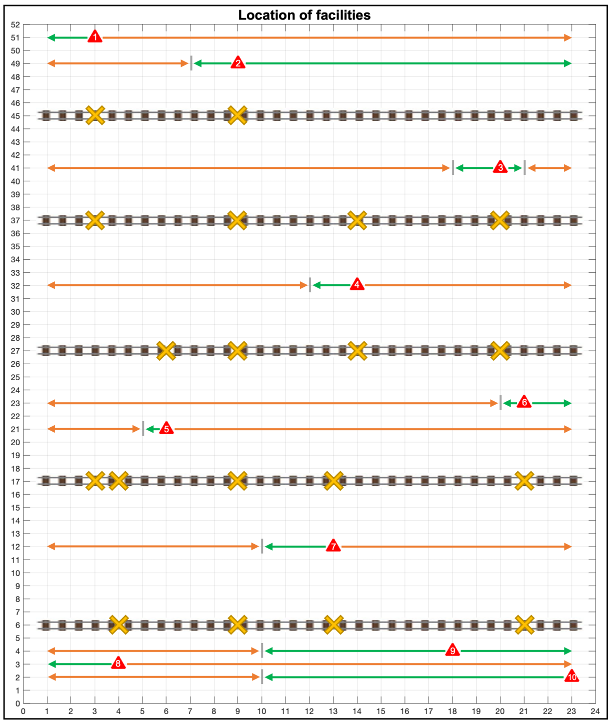
(a) Results for problem P01



(b) Results for problem P03

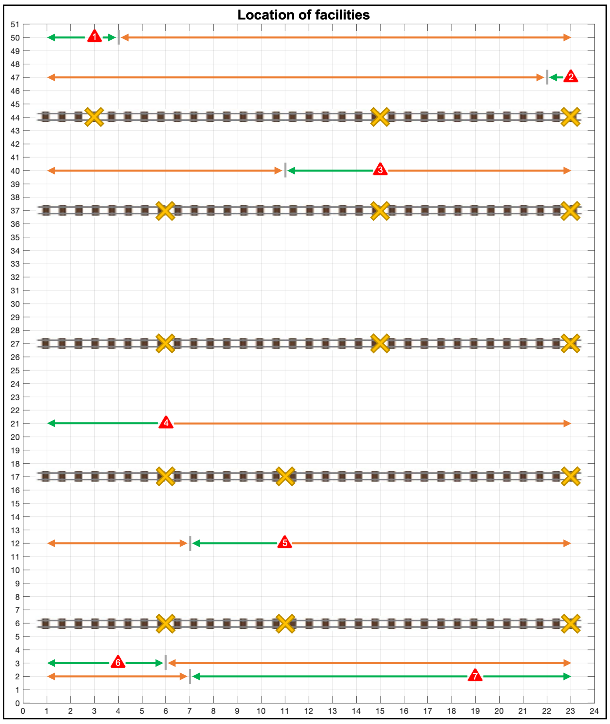


(c) Results for problem P04

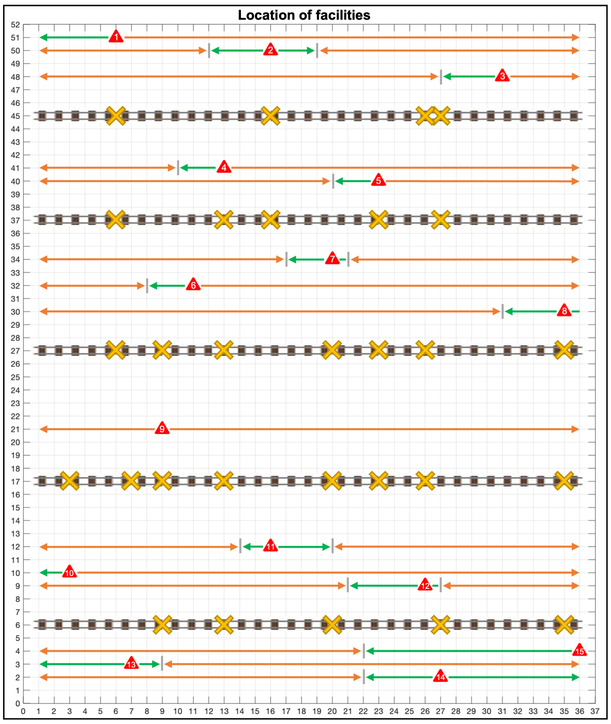


(d) Results for problem P05

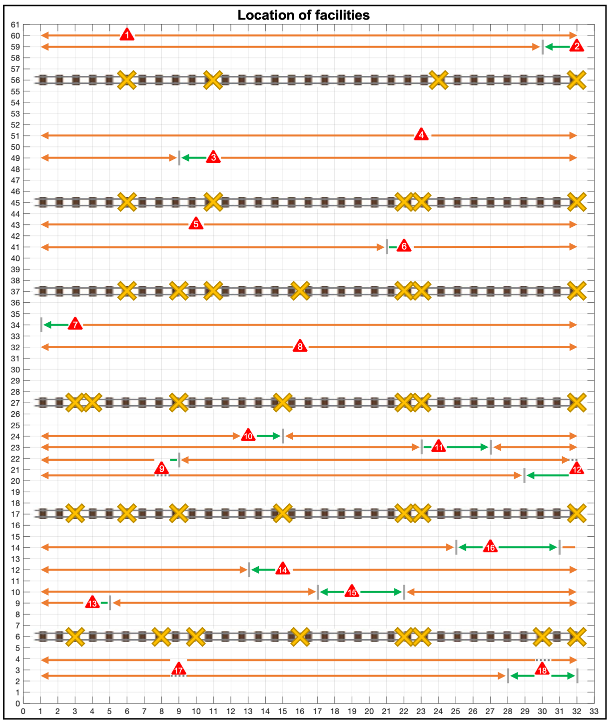
Appendix B. Results of sensitivity analysis for problems P06, P07, P09 and P10



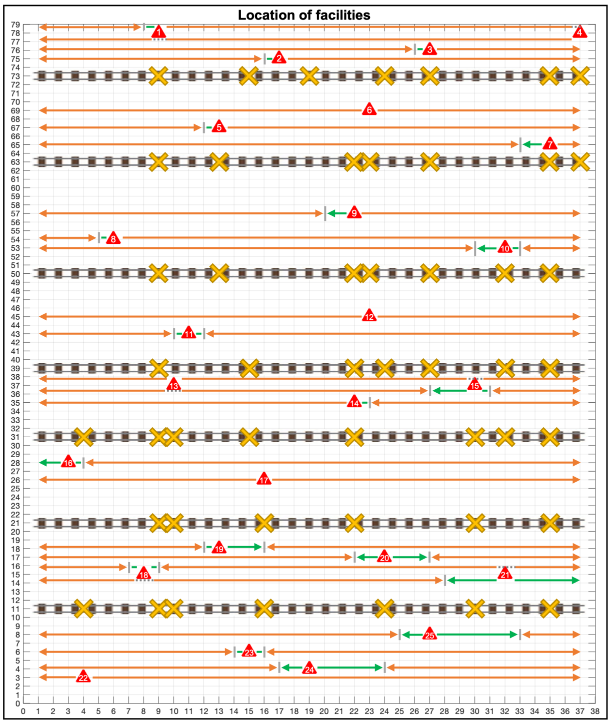
(a) Results for problem P06



(b) Results for problem P07



(c) Results for problem P09



(d) Results for problem P10