UNIVERSITYOF BIRMINGHAM

University of Birmingham Research at Birmingham

Sensitivity enriched multi-criterion decision making process for novel railway switches and crossings − a case study

Boghani, Hitesh; Ambur, Ramakrishnan; Blumenfeld Mendonca, Marcelo; Saade, Louis; Goodall, Roger M.; Ward, Christopher ; Plasek, Otto; Gofton, Neil; Morata, Miquel; Roberts, Clive; Dixon, Roger

[Link to publication on Research at Birmingham portal](#)

ORIGINAL PAPER

Sensitivity enriched multi-criterion decision making process for novel railway switches and crossings – a case study

Hitesh C. Boghani[1], Ramakrishnan Ambur[2]* , Marcelo Blumenfeld[2], Louis Saade[2], Roger M. Goodall[3], Christopher P. Ward[3], Otto Plášek[4], Neil Gofton[5], Miquel Morata[6], Clive Roberts[2] and Roger Dixon[2]

Abstract

Background: Despite their important role in railway operations, switches and crossings (S&C) have changed little since their conception over a century ago. It stands now that the existing designs for S&C are reaching their maximum point of incremental performance improvement, and only a radical redesign can overcome the constraints that current designs are imposing on railway network capacity. This paper describes the process of producing novel designs for next generation switches and crossings, as part of the S-CODE project.

Methods: Given the many aspects that govern a successful S&C design, it is critical to adopt multi criteria decision making (MCDM) processes to identify a specific solution for the next generation of switches and crossings. However, a common shortcoming of these methods is that their results can be heavily influenced by external factors, such as uncertainty in criterium weighting or bias of the evaluators, for example. This paper therefore proposes a process based on the Pugh Matrix method to reduce such biases by using sensitivity analysis to investigate them and improve the reliability of decision making.

Results: In this paper, we analysed the influences of three different external factors, measuring the sensitivity of ranking due to (a) weightings, (b) organisational and (c) discipline bias. The order of preference of the results was disturbed only to a minimum while small influences of bias were detected.

Conclusions: Through this case study, we believe that the paper demonstrates an effective case study for a quantitative process that can improve the reliability of decision making.

Keywords: Switches and crossings (S&C), Railway tracks, Multi-criterion decision making, Pugh matrix, Sensitivity analysis

1 Introduction

1.1 Switch and crossing (S&C) design

Switches and crossings (S&C) have been an integral part of the railway since their inception. They are an important part of railway operations, enabling railway vehicles to travel on various routes while maintaining a safe interface between wheels and tracks. On the other hand, due to the time required for actuation and interlocking of switches, as well as their physical limits to mechanical forces, S&C can limit the movement of trains, and thus establish a fine balance between operational flexibility and overall capacity.

Despite the general advancements in technology of the past 200 years, there have been only marginal changes or upgrades in S&C, such as introducing automation and improving wheel-rail interaction. The current designs suffer from several issues, such as dry slide chairs, fasteners/fittings failures, loose stretcher bars, and point operating equipment that are a single point of failure. In contrast, other subsystems such as railway communications, traction and power, and rolling stock have been continuously

* Correspondence: r.ambur@bham.ac.uk
[2]Birmingham Centre for Railway Research and Education, Department of Electronic, Electrical and Systems Engineering, University of Birmingham, Birmingham B15 2TT, UK
Full list of author information is available at the end of the article

revolutionised by technological innovations that have led to significant increases in operational capacity.

It seems that the current design of S&C is reaching the maximum performance that can be achieved by incremental improvements; current S&C design is one cause of important bottlenecks in contemporary railways. As explained by [6], the limitations of incremental innovations follow the law of diminishing returns where, after a certain point of maturity, greater engineering efforts are required to achieve increasingly small marginal enhancements in system performance. (Fig. 1).

This is particularly problematic because switches and crossings are becoming the bottleneck for railway capacity because of their physical limitations. Field data presented in [20] shows that breakdown at the S&C is directly linked to delays in the mainline rail network. Under changing conditions as the case with the strong digitalisation processes occurring in the sector, adaptation can only be achieved by a radical redesign leading to a completely new internal structure [23].

The potential benefits of a complete redesign are the underlying premise of the project that this paper reports. The S-CODE (Switch and Crossing Optimal Design Evaluation) project [28], funded by the European Commission, focused on identifying and developing radically different technological concepts for the next generation of railway track S&Cs. The aim was thereby to develop radically innovative concepts which could overcome the current limitations of S&Cs.

1.2 Multi-criteria decision making
Despite the clearly established goals and ambitions of the S-CODE project, the process of designing railway subsystems is usually complicated and it involves a multitude of options, constraints, standards, and wider implications. Furthermore, considering that the railways are large-scale

safety critical systems, decisions upon costs and reliability are paramount, and there is no 'one-size-fits-all' solution to address the idiosyncrasies of each line.

When there are many possible solutions for a problem, reaching consensus over the best choice requires appropriate tools. The field of Multi-Criteria Decision Making (MCDM), and the more specific Multi-Attribute Decision Making (MADM) is rich with methods that can simultaneously assess a large but finite array of criteria. These include Analytical Hierarchy Process (AHP), Quality Function Deployment (QFD), Pareto analysis, Elimination and Choice Expressing Reality (ELECTRE), and Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), among others [25]. Each of these techniques has varying attributes and some are more suitable to use at certain stages of projects than others.

Among them, the Pugh Matrix stands as a well-established tool [18] that is particularly helpful to evaluate different ideas during the early stages of a product's design, as it assesses the overall concept from a multitude of aspects. This method is particularly useful when a suitable product already exists as this then becomes a benchmark for other concepts to be evaluated. It is also simple and straightforward to use. Considering that the aim of the S-CODE was to look at potential overall designs for S&C, we adopted the Pugh Matrix as the method for our investigation. This matrix consists of a range of defined criteria such as novelty, ease of fabrication, and maintenance and production costs, across which new concepts are rated using a suitable scheme. The criteria in the Pugh Matrix are often rated against a datum concept (i.e., existing product/solution) which is considered "neutral". The criteria themselves are also weighted so that the overall score of a concept is the weighted sum of the given marks. All potential product
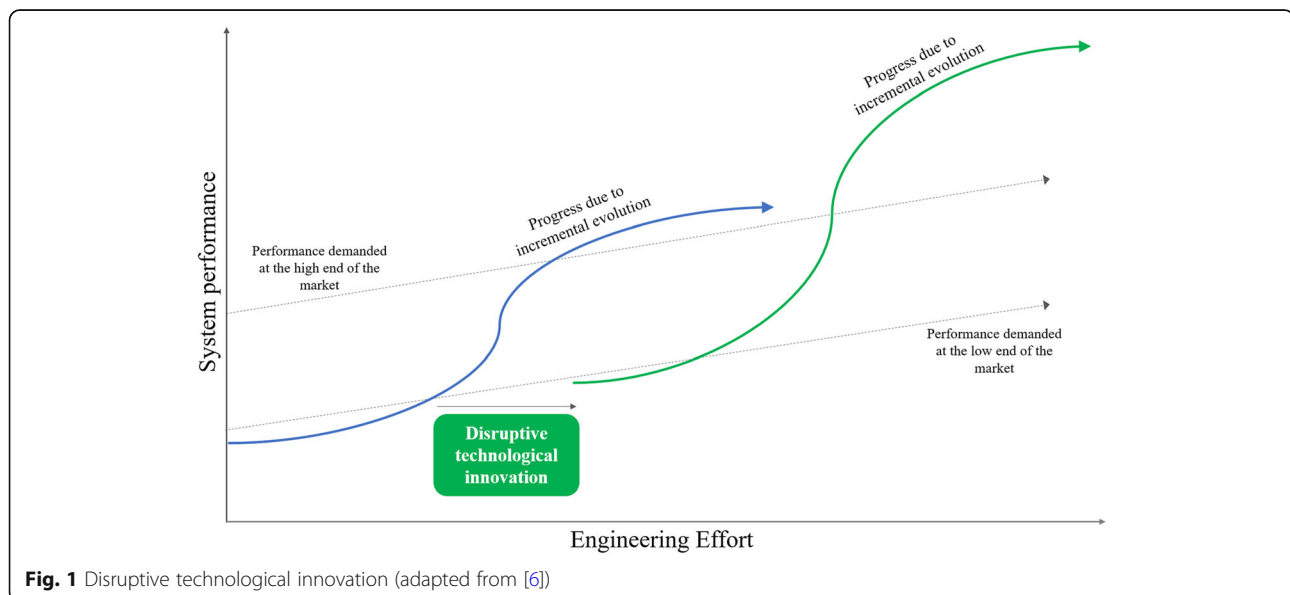

Fig. 1 Disruptive technological innovation (adapted from [6])

concepts can be ranked against one another based on the final weighted sum.

This concept selection process has been widely applied in product development. Recent notable examples include the design of an impulse turbine [24], selection of molecular instruments for use on the International Space Station [13], comparing cooling techniques for machining hardened steel [16] and carbonisation technologies for developing countries [14], design of a transformable chair [12], and selection of railway track switch concepts [2].

1.3 Sensitivity analysis

However, there are shortcomings attached to the Pugh Matrix method that require attention. An important example is that it does not lend itself to discriminating between the sensitivity of rankings to affecting factors. Ideally, external influences would not affect the rankings, but the reality is less precise in that sense. These external factors could be listed as uncertainty in weighting of the criteria, biased opinions of the evaluators due to their past experiences and concept favouritism. Even in small amounts, they can influence the weightings given and affect the reliability of the outcomes. Consequently, they can bias the criteria outputs for MCDM and lead to poor design choices.

This paper thus suggests a process to incorporate sensitivity into the evaluation from the Pugh matrix to reduce the influence of affecting factors as described above on the reliability of the outcomes. Sensitivity analysis on MCDM was considered by in one of the studies [26] but it was only focused on sensitivity due to the weighting of criteria. Also, they focused on finding critical criteria rather than looking at the distribution rankings given to a particular alternative concept by different evaluators. Goodridge [10] provided a technique to integrate sensitivity analysis to MCDM but it does not provide a probability distribution of the ranking and also does not consider biases from the assessors. By varying criteria weight one-at-a-time (OAT), one can obtain distribution of ranking received by particular alternative concept as a function of change in weight as used by Chen et al. [5] for GIS-based land suitability evaluation and was adopted here.

This paper describes the use of sensitivity enhanced Pugh Matrix to increase the reliability of ranking design concepts out of a list of potential candidates for decision-making processes. The various concepts and their evaluation process are explained in Sections 2 and 3, respectively. The suitable concepts resulting out of the evaluation are then ranked presented in Section 4. The susceptibility of these rankings to various biases in input data is verified by sensitivity analyses which is also presented in Section 4.

2 Methods

2.1 Identification of novel S&C concepts

Building on the project's ambition to develop radically innovative S&C systems, a consortium of project partners identified 22 high-level S&C concepts that aim to overcome the shortcomings of today's design, over brainstorming meetings. Some of the concepts are only applicable to a specific environment. These concepts are listed in Table 1, which are either new ideas or have some similarities or examples taken from S&C concepts being currently used in various parts of the world. This table is a consolidated list of concepts and ideas generated (90+) through literature search, project meetings and workshops. Consolidation was carried out based on variations within similar concepts, materials and, functional process used. Further reduction may be possible however, did not seem necessary at this stage. The complete description of all concepts is given in the S-CODE dissemination website [22].

2.2 Assessment of concepts and their evaluation

These high-level design concepts needed to be critically evaluated according to a defined criteria. This was intended to identify whether the potential concepts were to actually be beneficial compared to the conventional S&C or might worsen the situation. Hence concepts are subjected to the selection process using the Pugh matrix as presented in this section. The aim was to identify a set of concepts which are better than a benchmark, here a conventional S&C.

2.3 Criteria for assessing the concepts

The criteria against which the concepts have been evaluated are derived from [7] and then categorised further. This list is given in Table 2 along with the weightings for each. The criterion 'Radically different' was included not only due to the overall aims of the project, but also because of the systemic limitations of current designs. Current S&C systems operate with several actuation phases that, even if ultimately optimised, would only lead to marginal capacity gains. It logically follows that the innovative potential of each design should be considered in order to avoid optimisation solutions. The weightings imply the relative significance of a criterion amongst the set according to the method used in [2], as decided by a subgroup of five S-CODE consortium members.

Using their method, the weightings were derived by preparing a square matrix where criteria are arranged in rows as well as columns. The elements in this matrix were given values 0 if row criterion was less important than column, 0.5 if equally important and 1 if more important than column criterion. Diagonal elements were ignored as a criterion cannot be compared against itself. The total of all columns (row wise) was then divided by the grand total of all elements in the matrix to normalise total scores received by each criterion. This provided relative weighting of each criterion, to a total of 1 (i.e. 100%). The evaluation process ranks the concepts based

Table 1 List of devised concepts for novel S&C design

Concept	Description	Examples/Similarities
A	The track can be bent towards the appropriate route to be taken	(Samuel David Bemment, Dixon, & Goodall, 2013) [1]
B	Track direction can be changed or discontinuities eliminated by components which vertically move	(TU Delft, 2016) [27]
C	The whole switch section can be moved laterally	(Zimmermann, 2011) [30]
D	The whole switch section can be rotated around a horizontal axle	(Borras, 2011) [4]
E	The gap in crossing can be partially filled and wheels can be allowed to run on flange tips through this section	("Railway Systems- Frogs," 2019) [19]
F	The switch rail can be rotated to come into contact with the stock rail rather than by lateral movement	
G	The crossing nose can be actuated laterally to avoid transition gaps	("Frogs - Moveable Point Frogs (MPF)," 2019) [9]
H	Discontinuity in the crossing can be eliminated by moving wing rails	(Reinoehl & Henderson, 1909) [21]
I	Additional parts in track system to support the train while guiding wheelsets	(Wilson, Toht, & Katz, 1997) [29]
J	Panels can be moved to fill flangeway gaps with inserts	(Network Rail, 2017) [17]
K	Stock rail (or part of it) rotates around its longitudinal axis, independently from other stock rail	
L	Crossing could be substituted by a pivotable rail	
M	Flangeway gap in crossing nose could be eliminated by crossing rails at two different elevations	
N	Trains can be diverted from one rail route to another with a diverging bridge	(Network Rail, 2017) [17]
O	Steering can be done from bogie rather than track	(Farhat, Ward, Goodall, & Dixon, 2018) [8]
P	Track equipment can guide the back of flange of wheel	(Bezin, Kostovasilis, Neves, & Foan, 2018) [3]
Q	Guiding using flanges on wheels, on one side of bogie. The other wheel side need not have a flange.	(Network Rail, 2017) [17]
R	Many small sections of switch can be actuated to create a different route	(Kikuchi & Onaka, 1988) [11]
S	The wheel flange can be actively extracted to select the route and retracted to allow passage over a crossing section that has no gap	
T	The bogie can be guided with a single section of switch rail	
U	Dynamically filling gaps using deformable materials	
V	Mechanisms which actuate rails below the ground or hold them upright dependent on route to be taken by train	

on weighted sum of the marks received for each given criterion using weightings received by this process.

2.4 Data collection and processing

Members of the consortium were provided with a blank Pugh matrix, of which a partial version is shown in Table 3. In this matrix, each concept must be scored against each criterion on a scale between 0 and 10. Score 5 means that the criterion remains unaffected for that concept with respect to the benchmark. Scores lower than 5 and higher than 5 mean negative and positive contributions, respectively, to that concept compared to the benchmark, the conventional track S&C design in this case. Concepts that concerns novelty in only switch or crossing section were conceptually perceived to work with conventional counterparts. These were then compared against conventional S&C. This was to achieve comparability for all concepts that were included in this evaluation.

The Pugh matrix is pre-filled with a neutral mark, i.e. 5, against most criteria for the benchmark S&C (shown in italics in Table 2) except for criteria numbered 1, 3, 6 and 18 in Table 2, that cannot be marked as neutral (5) and are marked to be either 0 or 10 as the case may be. For example, criterion 1 for a conventional S&C must be marked as 0 because it cannot be radically different, and it has been functionally same design for a long history of railways and implemented worldwide already. Table 3 shows an extract from Pugh matrix, where candidate concepts are evaluated by an engineer.

Data for evaluation was collected from a wide group of people employed in eight railway specific organisations spread across four countries in the EU namely, Austria, Czech Republic, Spain and the UK. This

Table 2 Criteria used for Pugh Matrix assessment

Objective	#	Criteria	Description	Weight (%)
Design	1	Radically different	Is the concept/solution radically different to current/existing practice?	10.53
	2	Versatile and scaleable	Can the design be scaled up/down geometrically (such as length) and so is it versatile enough to be adapted to different types of track switch applications (speeds of operation)?	7.31
	3	Retrofitting	Is the solution retrofittable to the existing switching and crossing parts without many changes required in the existing system (track and vehicle)?	0.88
	4	Modularity	Is the design modular enough so that one can replace a faulty component with "off the shelf" components (or line replaceable units) without requiring to do any special manufacture?	7.6
	5	Allows track continuity	Does the design allow track continuity (i.e. no track gaps causing vertical loads)?	7.31
Manufacturing	6	Existing machinery / process can be used?	Can the outcome of the design be manufactured using existing production methods and machines without needing to develop a new manufacturing process/machinery?	1.46
Maintenance	7	Easy maintenance	Can the product be maintained easily? Could the maintenance be carried out with minimal labour, minimal machinery requirements and minimal track access time?	7.02
	8	Allows maintenance to be done offsite	Can the maintenance be performed offsite (i.e. replace parts on-site and repair the broken/failed part/component offsite)?	3.8
Logistics	9	Deployability	Can the product be deployed easily using minimal labour, minimal machinery and minimal track access time?	7.31
	10	Plug and play?	Does the product allow "plug and play" such that one can fit the parts and it is ready to use straight away?	7.31
Operation	11	Energy efficiency	Could the product be efficient in its energy use?	1.46
	12	Speed of switching	Can the switching operation be carried out in a quick enough time window?	4.39
	13	Improvement in loading?	Does the concept offer improvement in static and dynamic loading from existing designs?	5.56
	14	Weather resistance	Does the concept offer resistance to most adverse weather conditions such as extreme heat, snow or flooding?	6.72
Safety	15	Risk of derailment can be reduced	Can the concept reduce the risk of derailment?	8.19
	16	Allows safe run-through	Does the concept allow safe run-through when the train approaches from the wrong side of the switch?	3.22
	17	Reduction of out of correspondence	Is the concept robust enough to limit the "out of correspondence" (unsafe) state where the switch is not set to any route?	6.72
Other	18	Time to market	Does the concept require a minimal development effort before implementation?	1.46
	19	Cost	Is the solution likely to cost more than reasonably expected in comparison to the opportunity cost?	1.75

Table 3 An extract from Pugh matrix showing selected concepts and criteria

Objective	#		Conventional S&C	Concept A	Concept B	Concept C	Concept D	Concept E
Design	1	Radically different	0	8	9	7	8	5
	2	Versatile and scaleable	5	7	6	4	4	7
	3	Retrofitting	10	7	7	4	5	8
	4	Modularity	5	9	5	3	3	7
	5	Allows track continuity	5	8	10	7	7	9
Manufacturing	6	Existing machinery/process can be used?	10	5	5	5	5	5
Maintenance	7	Easy maintenance	5	7	5	4	4	7
	8	Allows maintenance to be done offsite	5	6	5	3	3	6

included academic researchers and engineers in industry, working on railway track related projects, as shown in Fig. 2. A total of 20 responses that were provided either by an individual or groups were gathered. The group evaluations were assumed to be an average assessment from the engineers within that group. The evaluators in the study are engineers affiliated to participating organisations who are railway regulators, permanent way constructers, S&C component manufacturers and system engineers. Given the cross-disciplinary nature of the system, categorising according to their technical skills and competencies as shown in Fig. 2 was considered a more suitable approach, rather than the nature of the organisations.

2.5 Evaluation and ranking

Each mark received by a concept against a criterion were averaged across all 20 responses and then weighted against corresponding weights given according to Table 2 and eq. (1).

$$M_w = \frac{W}{i} \sum_{i=1}^{n} m_i \qquad (1)$$

where M_w denotes weighted average mark (in contrast to M for average marks), W is given weight, m is mark, i is real number > 0 and, n is the number of set of evaluations, in this case, 20. Similarly, standard deviation of responses was also calculated to assess the degree of agreement across each concept. This was also multiplied with given weights

to focus only on the relevant variations according to the importance of criteria. This was obtained using eq. (2).

$$SD_w = w \sqrt{\frac{\sum_{i=1}^{n} (m_i - mean\ mark)^2}{n}} \qquad (2)$$

where SD_w is weighted standard deviation and other terms same as before. The weighted average marks were added together for each concept to obtain a weighted sum of all marks received by the same concept. These weighted sums (overall scores) were used to rank all concepts, including the benchmark, against each other.

2.6 Sensitivity analysis

Sensitivity analysis is a method for analysing the effect of uncertainty in the output of a system, subject to uncertainties in the inputs. To observe the effect of uncertainties, the inputs to the system are gradually varied and their corresponding effect on the outputs is studied. In the present case, the output of the evaluation study is the ranking of concepts. The inputs are the weightings of the criteria and the evaluators themselves. For each of these sources, two different sensitivity analyses were performed to identify any unconscious bias towards certain concepts.

In the first analysis, bias in evaluation from an organisation has been analysed. It is possible that one organisation favours a concept due to previous projects or expertise. Hence in the first analysis, evaluations from
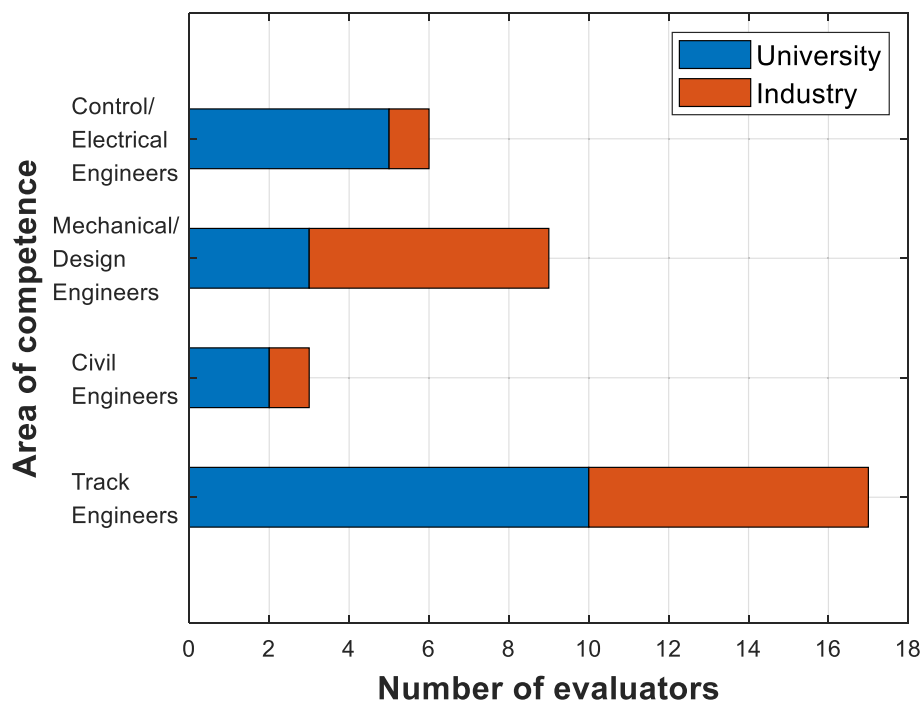

Fig. 2 Technical profile of evaluators showing diverse expertise

each organisation were removed one at a time and the concepts ranked again. The results are presented in section 3.2.1. Similarly, engineers sharing the same expertise could be biased towards a particular concept. This has been analysed in second sensitivity analysis in section 3.2.2.

At the outset the original weightings as given in Table 2 were agreed between four participants in the project. Because of this smaller expert group, the weightings as listed in Table 2, could be biased. To alleviate it, two more sensitivity analysis were performed. In the third analysis in section 3.2.3, two other experts calculated weightings for all the criteria. The concepts were evaluated again with these two new sets of weights prepared by separate experts.

In the fourth analysis presented in section 3.2.4, a holistic approach was developed where the weightings were sequentially changed. Each weight w_i is monotonically increased or decreased until it reaches ±100% of its initial weight, according to eq. (3). The subscripts a and g refer to the altered and given weights respectively. To maintain the sum of all weights as one, the weightings w_j of other criteria are adjusted to reflect this increment as described in eq. (4), where c represents the total number of criteria.

$$w_{ia} = w_{ig} + \Delta w_{ig}; for \left\{0 < w_{ia} < 1, \left|\Delta w_{i_g}\right| \le w_{i_g}\right\} \quad (3)$$

$$w_{j_a} = w_{j_g} - \left(\frac{\Delta w_{ig}}{c-1}\right); for \left\{0 < w_{j_a} < 1, j \neq i\right\} \quad (4)$$

The above-mentioned method leads to thousands of evaluations with its own rankings. Instead of comparing each outcome, the frequency of the rankings is then calculated as a percentage for each concept. The Monte Carlo method [15] uses random inputs to a system and looks at the outputs to deduce the probability with which the output changes. In the following method, instead of random inputs the weights were varied sequentially "one at a time". This is procedure to observe the sensitivity of rankings to the weights is also followed in [5].

3 Results

3.1 Ranking

Figure 3 shows the weighted average marks and weighted standard deviation for all criteria for the exemplary concepts shown in Table 3. The weighted average marks for the remaining concepts are presented in the annexure. The variation observed against all concepts for criteria 18 and 19 (see Table 2) are insignificant as shown by Fig. 3 since they are weighted relatively lower than criteria 1, 2, 4 and, 5. The converse can be observed from criterion 1, where because of its weighting, the variation observed is larger than unweighted (for example, ±1.8 for concept A, not shown). Similarly, though criteria 5 and 7 apply same weighting, the engineers have gauged them differently which explains their larger variation for criteria 5 than for
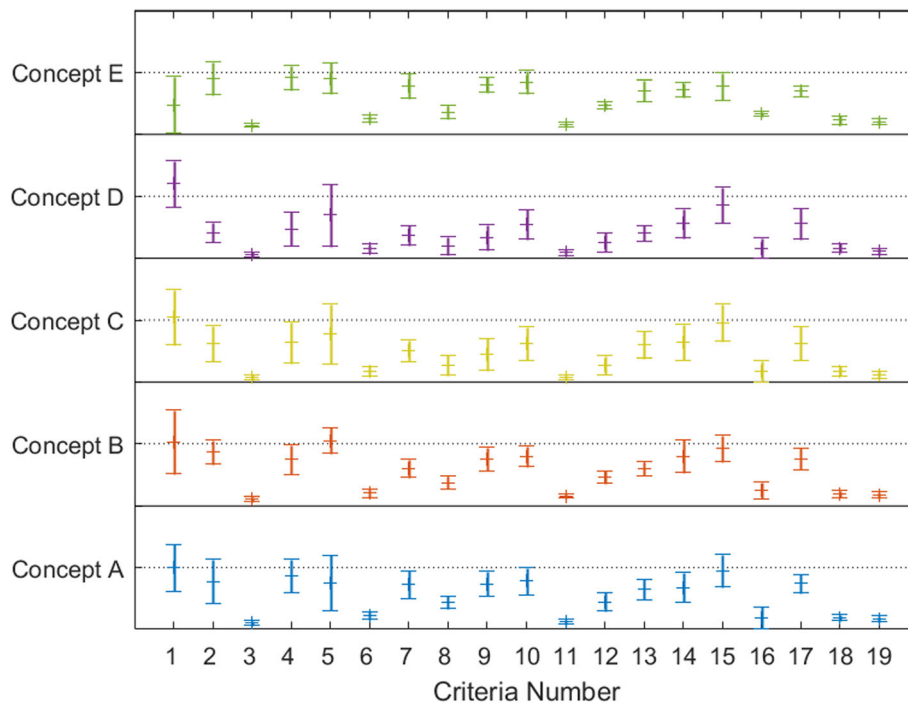

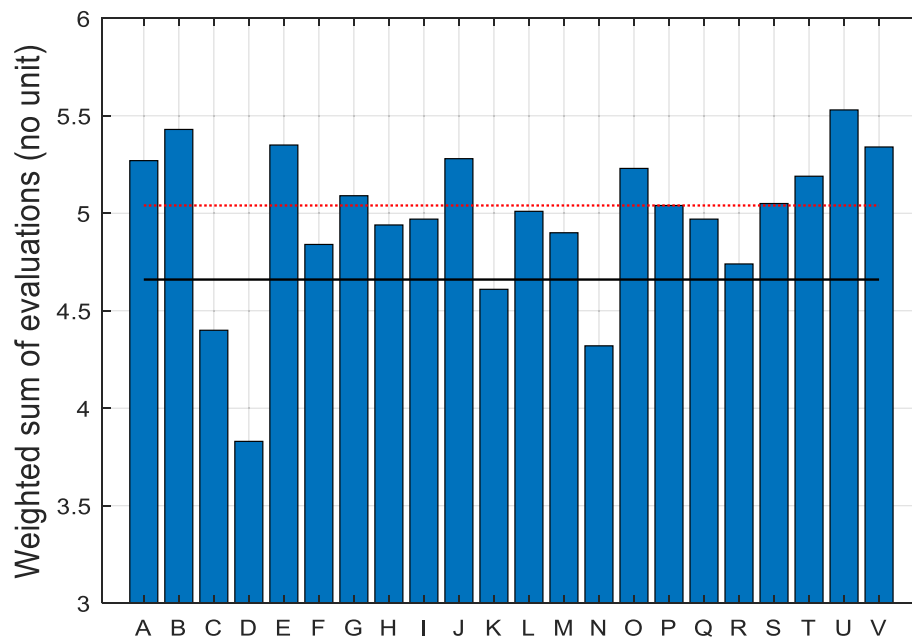**Fig. 3** Evaluation of selected concepts

Fig. 4 Weighted sum of evaluations from Pugh matrix

7. Criterion 5 concerns with 'allowing track continuity' and this can be subjected to the imagination of possible design solution around the concepts that were presented. So, it seems there was greater subjectivity seen by observing the standard deviation results for criterion 7. Reasons for large deviations in marking are:

- The explanation available was not sufficient to provide the same level of understanding amongst the evaluators.
- The experience of evaluators varied largely for those criteria; for example, manufacturing.

The weighted sum of average marks for each concept is compiled and shown in Fig. 4. The solid horizontal line indicates weighted sum for conventional S&C and dotted horizontal line demarcates the top 10 ranking concepts from others. The overall average marks received by each concept has a variation of − 23.4% to + 10.6%. The higher values indicate that concept is favoured compared to the benchmark concept. The top 10 most favoured concepts which can be read out from Fig. 4 are U, B, E, V, J, A, O, T, G and S in order of their rankings.
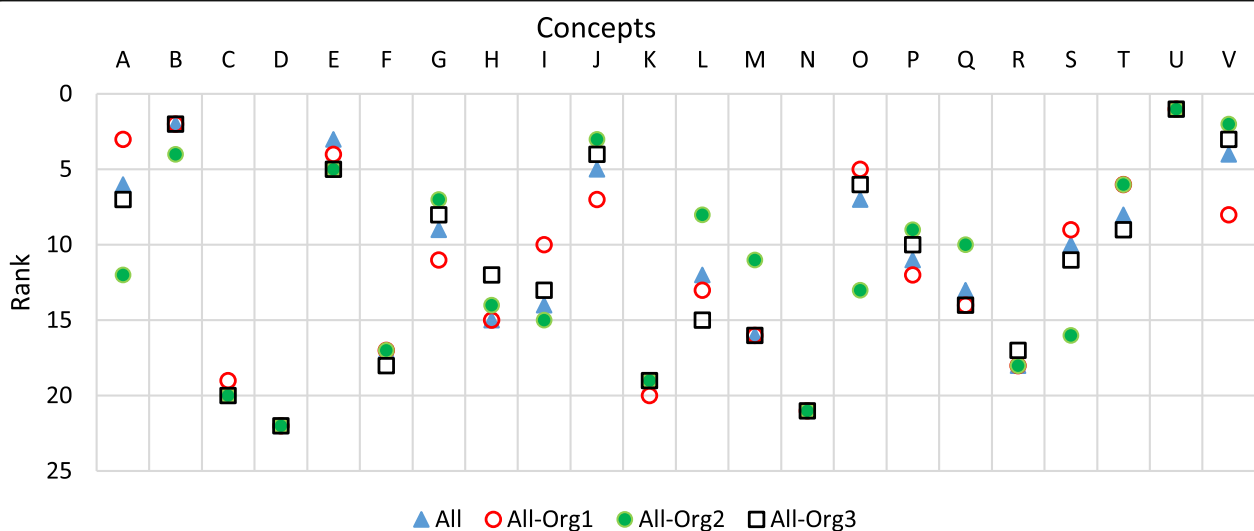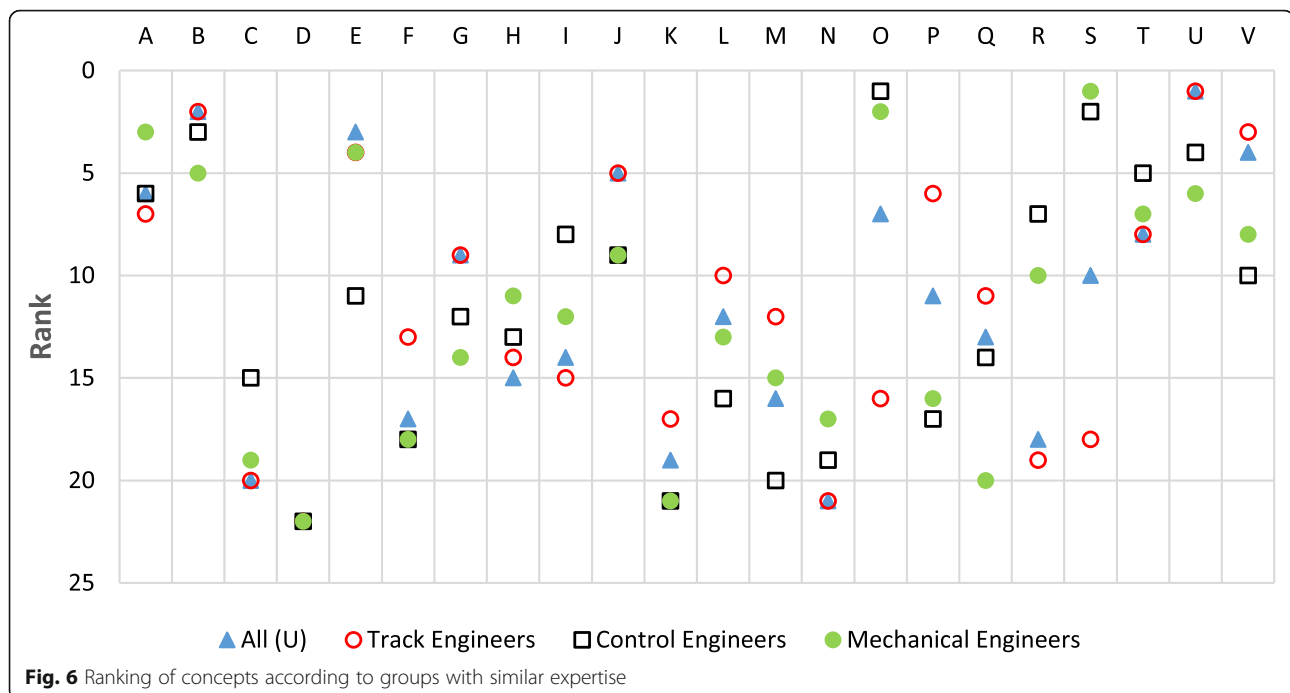

Fig. 5 Changes in rankings with evaluations of each organisation

Fig. 6 Ranking of concepts according to groups with similar expertise

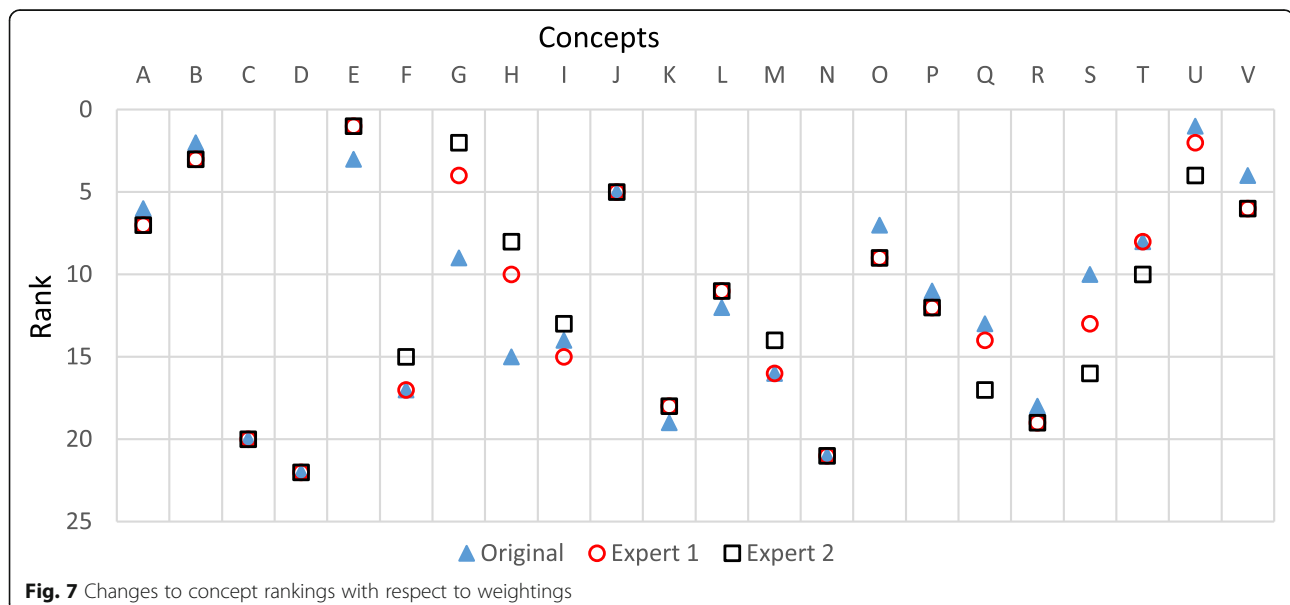## 3.2 Observations from sensitivity analysis

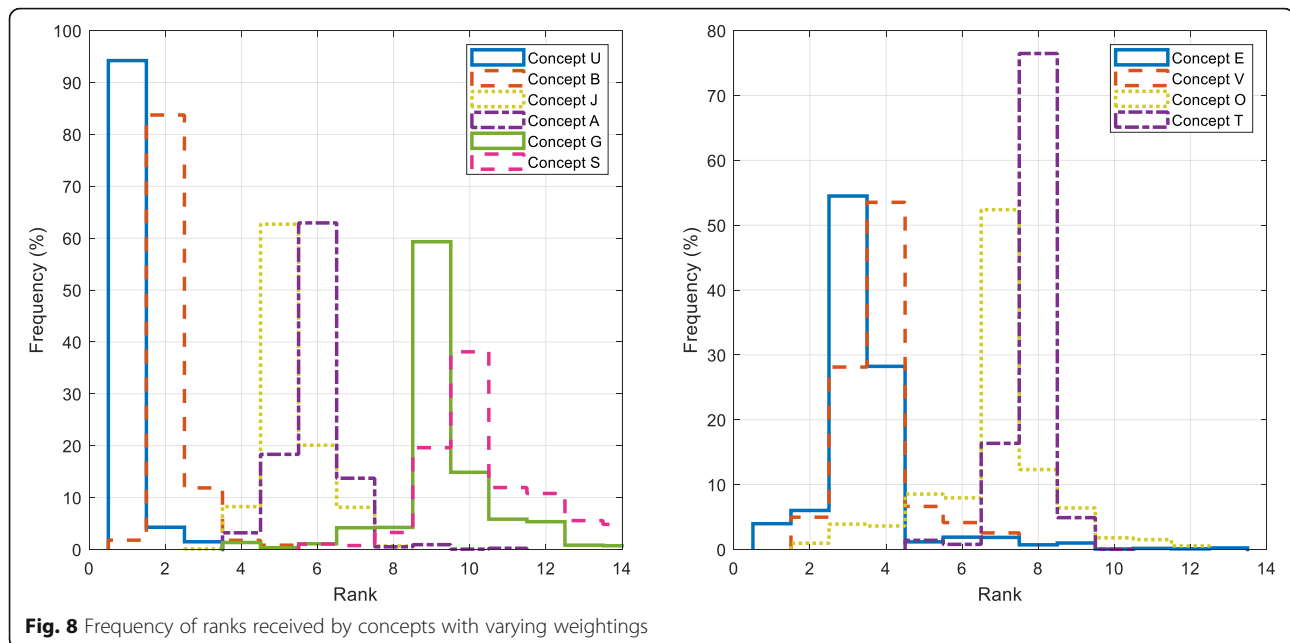### 3.2.1 Sensitivity to organisational influence

Figure 5 shows ranks of each concept by repeating the evaluation after one organisation was removed from the complete assessment at a time. It shows the complete evaluation as indicated by "All" and then evaluation when organisation 1 was removed as "All-Org1". For reasons of brevity, only four cases are depicted in Fig. 5.

As an example concept A can be analysed here. It was ranked to be sixth in the complete assessment. However when the assessment of organisation 1 was removed from the evaluation, its rank improved to 3. In similar lines, if organisations 2 and 3 were removed one after other from the complete evaluation, ranking of concept A dropped down to 12 and 7 respectively. This indicates that organisation 2 has favoured concept A through its evaluation.

Concepts C, D, F, K, N and R ranked consistently worse than 15th and hence are clearly not favoured by any party. Similarly concepts B, E, J, T, U and V ranked consistently better than 10th implying that there was consensus in selecting these concepts to take forward in the project.

Fig. 7 Changes to concept rankings with respect to weightings

Fig. 8 Frequency of ranks received by concepts with varying weightings

Certain concepts were indeed sensitive to group exclusion as observed from Fig. 5. The rankings for concepts A, O, S and V differed very much between their lowest and highest rankings. Concepts A, O and S scored poorly when ORG2 was removed from the assessment. Similarly, concept V was favoured by ORG1 as inferred from the evaluation. The disagreement observed from here may be due to a favoured concept due to the organisation's previous projects or expertise.

3.2.2 Sensitivity to expertise of evaluators

Figure 6 shows the ranking of concepts when the analysis was grouped based on expertise of the evaluators.

Whilst there was consensus on the less favoured concepts, there seemed to be a polarised view on some of them. Some important observations on the top 10 ranked concepts can be drawn from this analysis:

- Track engineers did not favour the concept 'O' which is based on vehicle active steering and rather preferred concept 'P' which is based on flange-back steering.
- Control engineers and mechanical engineers preferred concept 'O' (active steering) and concept 'S' (dynamic flanges), but not concept 'E' (flange bearing frogs) or 'V' (spring loaded pins).

These results show that the MCDM can be heavily influenced by the subject expertise and past experiences. Whilst this can be good in terms of informing the team members on realistic and achievable concepts, it could also hinder the creativity in terms of finding an elegant solution.

3.2.3 Sensitivity to consolidated weighting

As mentioned in section 4.1, Fig. 7 shows the ranking results by applying two separate weightings provided by two other participants from the project consortium. The results show that the rankings of the concepts listed are consistent despite different the changes to weightings by a different set of expert groups. If a demarcation line at rank 10 is considered, only the results from concept H and S are affected due to weightings. Because in other concepts, the ranks due to any weighting method, is either within or over rank 10.

3.2.4 Sensitivity analysis against individual criteria

To assess the sensitivity of ranking to individual criteria, the weightings were altered in a 'one at a time' fashion as described in Section 3.3. The frequency of the rankings from all iterations has been plotted in Fig. 8 for first five concepts out of the original evaluation as mentioned in section 3.3. It can be seen that concepts 'U', 'B' and 'E' are consistent in the rankings which indicate that they are least sensitive to changes in weightings. Concepts 'V' and 'J' are influenced with varying weightings, whereas concepts 'T' and 'G' showed tendency towards lower ranks (i.e. rank number increasing), but not shown for brevity as they are not ranked within the first five concepts.

4 Discussion

The basic MCDM analysis (Fig. 4) suggested top 10 ranking of concepts U, B, E, V, J, A, O, T, G and, S, rank 1–10, respectively. However, sensitivity analysis exposed the susceptibility of change in ranking to an evaluator and also to weighting of criteria. In particular, concept O and S showed large and equal spread of variations in their ranking around

their original (or base) rank as shown in Fig. 8. Also, their ranking variations due to evaluator's discipline (Fig. 6) was very large. This meant that these concepts did not have consensus in taking them forward within the project. Concepts U, B, E, V, J and A showed very low sensitivity to different weightings supplied by different partners within the project consortium which shows conformity in weightings applied to criteria.

Furthermore, sensitivity weighting of individual criteria tested by one-at-a-time approach showed that the concepts mostly swapped ranks between consecutive ranking concepts, for example, E and V and, J and A. Concept T and G had spread their ranking in opposite directions, T tending to score higher ranking (decreasing number) and G towards lower ranking. One-at-a-time sensitivity showed that although concepts U, B, E, V, J, A, T and G had variations in their ranking, mostly they all ranked within top 10 apart from concept G being at the edge (i.e. base rank 9) of the ranking range.

Additionally, the sensitivity analysis of section 3.2 enabled selecting the following concepts with confidence for further consideration.

- Switch only concepts: B, A.
- Crossing only concepts: U, E, V, J, and G.
- Switch and crossing concept: T

These subsets of concepts could be further explored with engineering methods for their feasibility to replace the conventional S&C. Upon looking at the root causes of the sensitivities to a particular criterion in these concepts, it is possible to improve concepts by extracting some features from other concepts and including into the original concept. For example, it may be possible to combine concepts 'G' and 'V' and create a novel concept such as spring loaded actuated nose, to receive the benefits of both and also overcome shortcomings of either of them, though it is not always possible to extend to all combination of concepts.

Using one-at-a-time analysis alone did not allow to show huge inconsistencies received by concepts O and S. This is important to know when a large team effort is required and the team here is formed of individuals from different organisations working towards a common goal. Inconsistencies in ranking suggests differing views and confidence on those concepts. By applying 'one-at-a-time' sensitivity tests on weightings and analysing those along with applying sensitivity through exclusion criteria allows one to spot the inconsistencies in the ranking and select the best candidate concepts with confidence as shown by the case study here.

5 Conclusions

This article has presented a process to incorporate sensitivity analysis into multi-criteria decision making methods (MCDM) through a case study that looked at novel designs for switches and crossings (S&C). Under the premise of the need for radically novel S&C designs, a panel of engineers across different domains of the railway sector were asked to evaluate existing concepts and novel ideas using a comprehensive set of criteria. The concept down-selection was conducted using the Pugh Matrix as the method for multi criteria decision making, which evaluates the concepts via marks given to them across the criteria with their weightings.

As hypothesised, the study showed that there are some biases in the weightings and results achieved with multi-criteria decision-making processes. These can be amplified when a panel of more than 20 engineers from 9 organisations with different fields of expertise. Despite having an overarching high average ranking, some design options can carry large deviations that demonstrate a significantly different perception of its overall quality or feasibility. This highlights the impacts of subjective perceptions that derive from professional expertise or institutional focus, that may have substantial influence on processes with small cohorts. This also raises the question whether it is better to seek homogeneous cohorts for cohesion, which on the other hand may threaten the innovative potential of the group.

The sensitivity of results to four various influencing factors was examined. In these analyses, the obtained results were checked for possible biases from engineers of different organisation, variation due to weightings and for preferential treatment by engineers of a particular expertise. A group of concepts was identified as more promising after being consistently ranked highly despite the influence of the above variations. The best ranked concepts were selected, which could be developed further for a railway track S&C of the future. The article shows that the added sensitivity analysis can highlight the impact of external factors on criteria evaluation, and with that improve the confidence of decision making to identify the best candidates out of the set. From the biases observed in different iterations, we can conclude that the sensitivity analyses can enhance the use of MCDM to choose amongst equally possible options.

Given the fast pace with which the engineering context is advancing, it is expected that novel designs will be required more often. At a time when railway systems become more complex and projects include a large number of experts working together, we anticipate the process hereby presented as a step forward to improving decision making for new solutions. It is important that personal biases from technical experience be identified and optimal choices are as free as possible from personal biases. However, in practice this is not straightforward. So, this paper has sought to incorporate the variation due to such biases, so that they can be taken account of in the analysis in order to make better decisions.
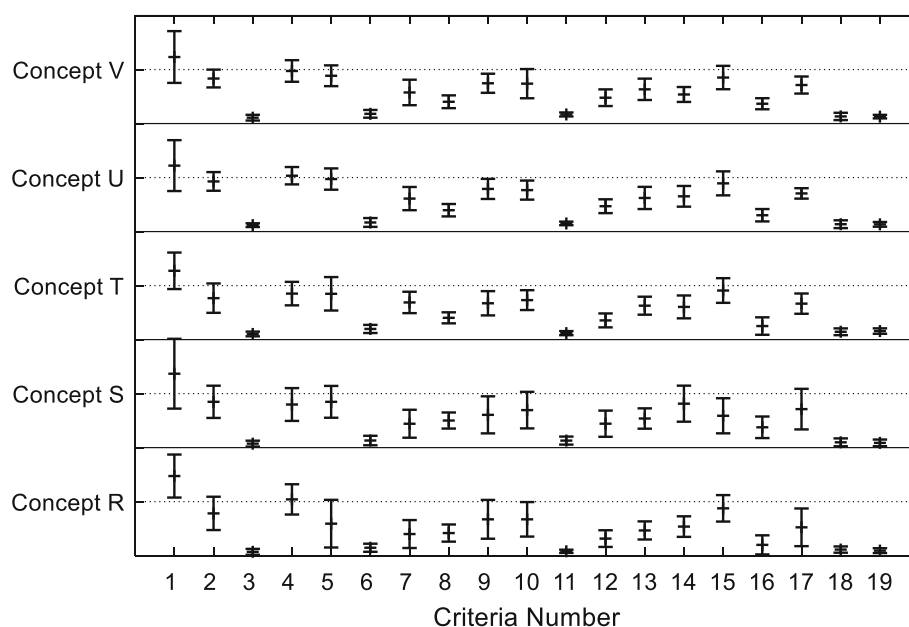
6 Appendix

Fig. 9 Evaluation of concepts F to K

Fig. 10 Evaluation of concepts L to Q
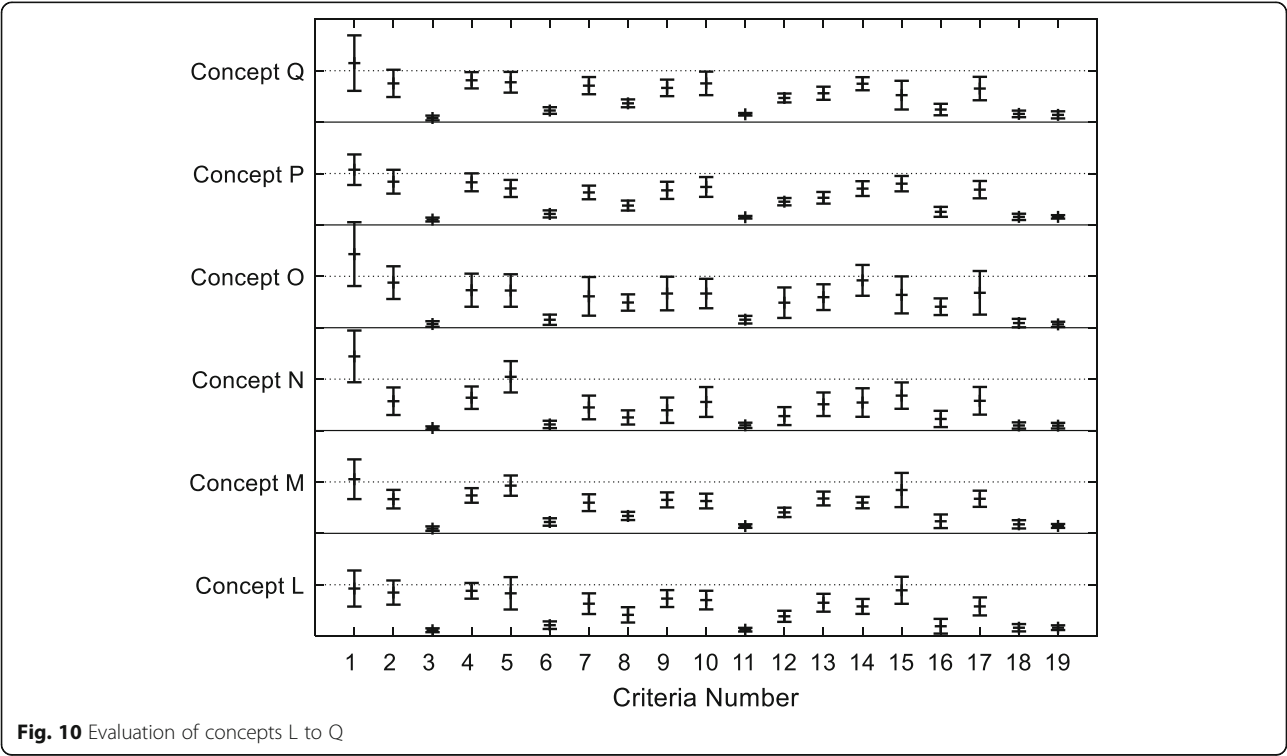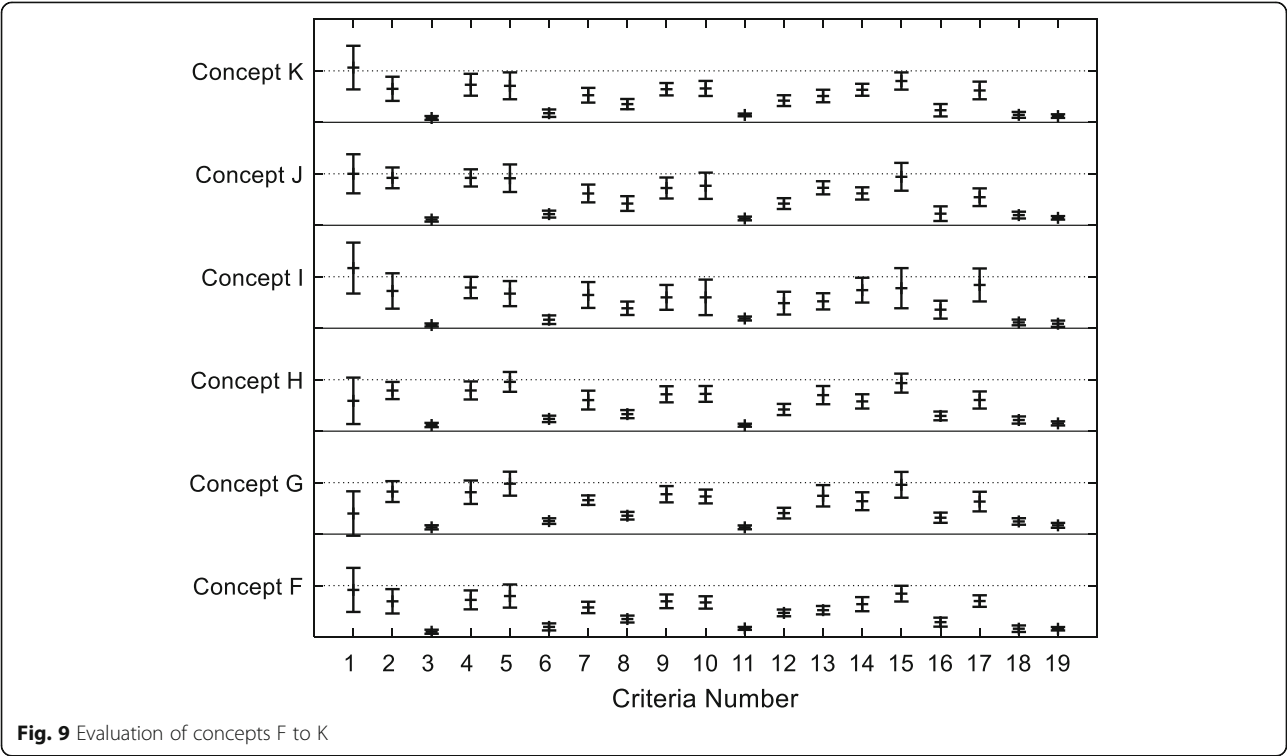
Fig. 11 Evaluation of concepts R to V

Abbreviations

S&C: Switches and crossings; S-CODE: Switch and Crossing Optimal Design Evaluation; MCDM: Multi-Criteria Decision-Making; MADM: Multi-Attribute Decision Making; AHP: Analytical Hierarchy Process; QFD: Quality Function Deployment; ELECTRE: Elimination and Choice Expressing Reality; TOPSIS: Technique for Order of Preference by Similarity to Ideal Solution; OAT: One-at-a-time

Authors' contributions

HCB, RD, MB, LS and CPW collected and consolidated the concepts data from different brainstorming sessions and defined the multi-criteria-decision-matrix (MCDM) template. HCB collected data for multi-criteria decision-matrix exercise, proposed sensitivity analyses criteria and HCB and RA processed the data. RA led the article writing including co-ordinating with co-authors. OP, NG and MM contributed to the study by developing independent weightings to apply within MCDM for sensitivity analysis against the benchmark. RMG, CPW, CR and RD supervised the entire technical process. All authors read and approved the final manuscript.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Transport Safety Research Centre, School of Design and Creative Arts, Loughborough University, Loughborough LE11 3TU, UK. [2]Birmingham Centre for Railway Research and Education, Department of Electronic, Electrical and Systems Engineering, University of Birmingham, Birmingham B15 2TT, UK. [3]Control Systems Group, Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University, Loughborough LE11 3TU, UK. [4]Faculty of Civil Engineering, Institute of Railway Structures and Constructions, Brno University of Technology, Veveří 331/95, 602 00 Brno, Czech Republic. [5]RSSB, The Helicon, 1 South Place, London EC2M 2RB, UK. [6]COMSA, Av. Roma 25-27, 08029 Barcelona, Spain.

References

1. Bemment, S. D., Dixon, R., & Goodall, R. M. (2013). *UK Patent No. GB2516707A*.
2. Bemment, S. D., Ebinger, E., Goodall, R. M., Ward, C. P., & Dixon, R. (2017). Rethinking rail track switches for fault tolerance and enhanced performance. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *231*(9), 1048–1065. https://doi.org/10.1177/0954409716645630.
3. Bezin, Y., Kostovasilis, D., Neves, S. G. M. D., & Foan, A. (2018). *Design evaluation of flange back steering for railway switches*. In *Paper presented at the the fourth international conference on railway technology, Sitges, Spain*.
4. Borras, B. (2011). Pilatus rack railway. Retrieved from http://www.trenscat.com/internacional/ch-pb_en.html
5. Chen, Y., Yu, J., & Khan, S. (2010). Spatial sensitivity analysis of multi-criteria weights in GIS-based land suitability evaluation. *Environmental Modelling and Software*, *25*(12), 1582–1591 https://doi.org/10.1016/j.envsoft.2010.06.001.
6. Christensen, C. M. (2000). *The Innovator's Dilema: When new technologies cause great firms to fail*. Boston: Harvard Business Review Press.
7. European Commission. (2015). Multi-annual action plan. Retrieved from Brussels:
8. Farhat, N., Ward, C. P., Goodall, R. M., & Dixon, R. (2018). The benefits of mechatronically guided railway vehicles: A multi-body physics simulation study. *Mechatronics*, *51*, 115–126.
9. Frogs - Moveable Point Frogs (MPF). (2019). Retrieved from https://www.voestalpine.com/nortrak/en/products/Frogs-Moveable-Point-Frogs-MPF/. Accessed 2 Feb 2020.
10. Goodridge, W. S. (2016). Sensitivity analysis using simple additive weighting method. *International Journal of Intelligent Systems and Applications*, *8*(5), 27.
11. Kikuchi, S., & Onaka, A. (1988). Monorail development and application in Japan. *Journal of Advanced Transportation*, *22*, 17–38.

12. Kulkarni, S. N., Reddy, J. K., & Gandhi, P. P. (2015). Conceptual development of a transformable chair using TRIZ. *International Journal on Design & Manufacturing Technologies*, *9*(1), 29–37.

13. Lineberger, K., Levitt, J., Smith, D. J., Van Nguyen, T., & Peter, A. M. (2014). A systems engineering approach to quantitative comparison of molecular instruments for use on the international Space Station. *Procedia Computer Science*, *28*, 340–346. https://doi.org/10.1016/j.procs.2014.03.042.

14. Lohri, C. R., Rajabu, H. M., Sweeney, D. J., & Zurbrügg, C. (2016). Char fuel production in developing countries – A review of urban biowaste carbonization. *Renewable and Sustainable Energy Reviews*, *59*, 1514–1530. https://doi.org/10.1016/j.rser.2016.01.088.

15. Metropolis, N., & Ulam, S. (1949). The Monte Carlo method. *Journal of the American Statistical Association*, *44*(247), 335–341. https://doi.org/10.2307/2280232.

16. Mia, M., Gupta, M. K., Singh, G., Królczyk, G., & Pimenov, D. Y. (2018). An approach to cleaner production for machining hardened steel using different cooling-lubrication conditions. *Journal of Cleaner Production*, *187*, 1069–1081. https://doi.org/10.1016/j.jclepro.2018.03.279.

17. Network Rail. (2017). In2Rail Deliverable D2.5 Radical S&C concept: Design concept evaluation study report. Retrieved from Milton Keynes

18. Pugh, S. (1995). Design Core: Conceptual Design. In *Total Design*, (pp. 67–99). Cornwall: Addison-Wesley Publishers Ltd..

19. Railway Systems- Frogs. (2019). Frogs - Partial Flange Bearing Frogs. Retrieved from https://www.voestalpine.com/nortrak/en/products/Frogs-Partial-Flange-Bearing-Frogs/

20. Rama, D., & Andrews, J. D. (2013). A reliability analysis of railway switches. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *227*(4), 344–363.

21. Reinoehl, C. W., & Henderson, W. M. (1909). *Steelton*.

22. S-CODE Consortium. (2017). D1.1 Review of definition, standard operating parameters, best practice and requirements, including future technologies and horizon scanning. Retrieved from Birmingham

23. Saviotti, P. P. (1986). Systems theory and technological change. *Futures*, *18*(6), 773–786.

24. Thakker, A., Jarvis, J., Buggy, M., & Sahed, A. (2009). 3D CAD conceptual Design of the Next-Generation Impulse Turbine using the Pugh decision-matrix. *Materials & Design*, *30*(7), 2676–2684. https://doi.org/10.1016/j.matdes.2008.10.011.

25. Triantaphyllou, E. (2000). *Multi-Criteria Decision Making Methods*. Berlin: Springer Science, Media, and Business.

26. Triantaphyllou, E., & Sánchez, A. (1997). A sensitivity analysis approach for some deterministic multi-criteria decision-making methods. *Decision Sciences*, *28*(1), 151–194.

27. Delft, T. U. (2016). WIRAS Winterproof- without point heating. In *Winterproof railway turnout* Retrieved from http://www.winterproofturnout.info.

28. University of Birmingham. (2017). S-Code - Switch and Crossing Optimal Design and Evaluation. Retrieved from http://www.s-code.info/

29. Wilson, R. V., Toht, D. E., & Katz, J. H. (1997). *US Patent No. US5826835A*.

30. Zimmermann, K. (2011, May). A new moon over New Hampshire. In *Trains Magazine*, (pp. 34–41).

Publisher's Note