A Collection of Aspects Why Optimization Projects for Railway Companies Could Risk Not to Succeed – A Multi-Perspective Approach

Christian Liebchen [a], Hanno Schülldorf [b]
[a] TH Wildau, Engineering and Natural Sciences
Hochschulring 1, 15745 Wildau, Germany
[1] E-mail: liebchen@th-wildau.de, Phone: +49 (0) 3375 508 755
[b] DB Analytics – Optimization
Poststraße 20, 60329 Frankfurt am Main, Germany

Abstract
You might be aware of the following gap: There are by far more publications on promising projects on how mathematical optimization could improve the performance of railway companies, than true success stories in the sense that operations research methods really entered the practice of railways.

In this paper, we shed a bit of light on those projects, which finally did *not* enter the practice of railways. We do so by conducting a survey in which we ask both, railway practitioners who served as ordering party, and optimization experts who served as R&D solution provider.

We summarize and comment the most frequent replies to our question about the key factors why in the past mathematical optimization methods did not enter the practice of railways: expert capacity for validation, management attention, quality of input data, and "moving target" objectives. Hereby, we offer a knowledge base to future project managers. Acting accordingly with respect to definition of project goals, project design, and project management, hopefully lets them come up with even more true success stories of operations research methods in the practice of railways.

Keywords
Railway Optimization, Operations Research, Project Management, Limiting Factors, Do's and Don't's

1 Introduction

Planning for and operations of railway systems are prominent fields of application for mathematical optimization models and algorithms. In particular, during the last decades there had been reported many projects in which in particular (mixed) integer linear programming technologies had been the technology of choice to solve the real-world problems of railway companies. The particular tasks to be covered include for instance:

- network design for cargo traffic
- line planning in passenger transport
- design of the basic hourly patterns for periodic timetables
- track allocation as it is usually performed by infrastructure managers

- vehicle scheduling (locomotives, passenger train units; rotations)
- crew scheduling (train drivers and/or conductors)
- shunting planning at shunting yards
- delay management (from an infrastructure manager's perspective and/or with the focus on passenger connections)
- schedule for ticket control staff
- and many more.

Some of these elementary tasks already have been even equipped with features which are reasonable to add, but which make the mathematical task itself even more complex. Think of periodic timetabling, for which there have been included robustness issues and demand feedback loops, in particular passenger re-routing. Due to the sensitivity of the subject that we are presenting in this paper, we refrain from providing references to particular papers. Rather, we generally refer to the references that are included in survey papers such as Borndörfer et al. (2018), Cacchiani and Toth (2012), Caimi et al. (2017), and Harrod and Gorman (2011).

Nevertheless, there do not seem to be dozens of papers that do not just report on promising projects, but rather on success stories in that operations research methods really entered the practice of railways. The most striking one is of course Kroon et al. (2009), for which the authors received the INFORMS Edelman Award – "The Oscar® of O.R." – for their various mathematical contributions for the Dutch Railways. Besides, other papers in which optimization results have been used and applied for railway operations, include for instance Kohl (2003) and Liebchen (2008). In contrast, we are only aware of the paper by Gorman (2016), in which he explicitly describes the failure of a railway optimization project.

To summarize, we feel there is a kind of gap between the number of projects in which mathematical optimization experts and railway practitioners work together, and the number of success stories in the sense that the operations research methods are applied on a regular basis in the practice of railway companies. This impression is based on our personal experience in a couple of projects on both sides: the practice of railways as ordering party, as well as research institutions as solution providers. Notice that we are not limiting ourselves to daily operations, but we would also consider it as a success story, if for regular strategic questions (such as in the context of public tenders) the respective methods are applied regularly.

This is why in this paper we shed a bit of light on those projects, which did not become a "success story" in the above sense. We are interested in such projects, whose project goal in the beginning has been the application of the developed mathematical optimization methods on a regular basis, but which did not attain this goal: Are there any common key properties, which prevented several of these projects to become true success stories?

We are aware of some personal summaries and collections of general hints on selected specific success factors for railway optimization projects, provided by some experienced railway optimization experts, e.g. Borndörfer et al. (2017) and Schülldorf (2018). Yet, we think it might be of interest to set such collections on a broader basis, both for the number of experts who are sharing their experience, and for the fact that both sides – including railway practitioners as ordering party – shall contribute with their experience. This is why we initiate a survey in which we investigate this question by asking both railway practitioners and mathematical optimization experts who were involved in such projects through a questionnaire.

The objective of this paper is to enable future project managers to setup their project goals and project design & management upon the negative experiences that other projects faced in the past. Hereby, we hope to improve the possibility that in the future more mathematical optimization projects for railway companies will become true success stories.

The paper is structured as follows. The questions that we are asking about the projects can be separated into two more or less separate classes. First, there is the general – and mainly administrative – framework of the projects (e.g., its duration, its partners, and its funding), which we describe in Section 2.1. Next, in Section 2.2 we add some problem-specific properties of the projects, some of which we suppose to be critical for a project to become a true success story. In Section 3, we shortly sketch the realization of our survey, before in Section 4, we present the results of our survey. On the one hand, we are aware that a number of 24 filled questionnaires is indeed "limited", and in particular far from being representative. On the other hand, to the best of our knowledge, this still constitutes the largest knowledge base in this specific field.

We report the results separately for the replies that we obtain from railway practitioners (N=10), and for the mathematical optimization experts (N=14), because there was a slightly different *awareness* regarding the most important reasons for project failures. Finally, we propose some conclusions for the future design and implementation of optimization projects for railway companies in Section 5.

2 General Administrative Framework of the Projects

In this section we essentially list the questions that we ask the former project members. We start with some questions to classify the projects according to some rather general properties in Subsection 2.1. Hereafter, in Subsection 2.2, we list our questions regarding rather content-related and method-related features that a project could show, and of which we can imagine that some of them might have significant influence why certain methods finally are not used in practice on a regular basis.

This distinction between the sets of questions is motivated by our goal to relate certain specific reasons for failure to some general framework properties of the projects (e.g. project duration), see Subsection 4.4 for some selected correlations.

2.1 General Administrative Framework of the Projects

In the sequel, we list the general properties of a project for which we ask in our questionnaire.

(a) Is it a railway practitioner of a mathematical optimization expert who is answering?

(b) Goal: Has the project goal been the reduction of (operational) cost and/or some increase of quality?

(c) Cost components: If the project goal was mainly cost efficiency, did the calculation of the estimated benefit of the project only include the expenses for the research, or also the full integration into the software landscape of the company including interfaces, education etc.?

(d) Funding: To what extent have the expenses for research been funded apart from public money from some research agency?

(e) Suppliers: Who has been responsible for the R&D part: universities, research institutions, software companies?

(f) Changes: Has the project goal been modified significantly during the project ("moving target")?

(g) Horizon: Which attribute fits best the target of the project: strategic decisions, planning for the operations process, the operations process itself, or any other?

(h) Target: Have the following been affected by the by the optimization results: vehicles, operational staff (train drivers, shunting assistants, conductors)?

(i) Timeline: What has been the timeline duration of the R&D part of the project (up to 12 months, 13-24 months, at least 25 months)?

(j) Volume: What has been the project volume of the R&D part in the sense of manpower (up to 12 months times men, 13-36 months times men, at least 37 months times men)?

(k) Urgency: Have there been alternative ways (without mathematical optimization) to come up with *some* solution(s) for the questions that should be answered by the new optimization methods?

(l) Input: What has been the structure of the input that was necessary to feed the mathematical optimization models: all data – except for optimization specific parameters – have been available in one existing IT system, all data have been available in IT systems but had to be combined from more than just one system, some of the data that had been necessary to feed the mathematical optimization models had not been available in any existing IT system?

(m) Output: What has been the existing IT-infrastructure to receive and further process the result of the mathematical optimization: Does the optimization result have the same data structure as it is already stored in some IT system(s), e.g., to manage solutions that earlier had been designed manually, or is there any manual post-processing required to fit the optimization data into the existing IT-infrastructure, or is there even a completely new IT system or organizational structure required in order to further work with the optimization results?

(n) Interpretability: How complex is it to "understand" the solution returned by the mathematical optimization model? It is just accessible at the level of key performance indicators (KPIs), or is it possible to comfortably dive into the very details of the solution, maybe even supported by some appropriate visualization?

In addition to these rather organizational properties of optimization R&D projects for railway companies, in the next section we present rather content-related and method-related features that a project could show, and of which we can imagine that some of them might have been decisive for the lack of success of some particular projects.

2.2 Problem-Specific Properties of the Projects

Now, we switch from the organizational perspective on the projects' framework to some of their content-related and method-related properties, which seem to us to have the potential having been a limiting factor more than only just one time. In the summary of the results of our survey (Section 4), we will put emphasis on these features, in order to identify those constellations which in the past had been the most likely show-stoppers for optimization projects for railway companies to become true success stories – and on which future projects should pay most attention right from the very beginning.

(1) Data: The available input data finally did not meet the quality that was necessary to be able to come up with high-quality optimization results (e.g. better than solutions that were designed manually).

(2) Partial Fixing: The optimization missed the ability to accept some particular fixation for certain “variables” that were key in the point of view of the railway practitioners.

(3) Features: During the project timeline, the optimization model had been confronted with more and more detailed requirements, which finally let the performance and/or quality of the optimization methods collapse.

(4) Validation: The railway company didn’t allocate a sufficient amount of expert staff to validate in detail the results of the optimization methods during the entire project timeline.

(5) Post-processing: The optimization environment lacked an editor that enabled the railway practitioners to (slightly) adjust the solution that was returned by the optimization algorithm to meet their actual practical needs and expectations?

(6) Quality: The optimization results failed to outperform the previously manually designed solutions and/or the optimization results did not achieve the quality which has been assumed in the cost-benefit-analysis that had been the basis to initiate the project.

(7) Regularity: The optimization results didn’t show a certain “regularity pattern”, which in the end had been expected by the railway practitioners (although not communicated as a key feature at the project kick-off).

(8) Transparency: The structure of the optimized solutions stayed somehow intransparent – “sealed” – to the railway practitioners which let them refrain from continuing to work with them in the sequel.

(9) Integration: The solution indeed optimized the specified task, but from a process perspective, subsequent tasks let expect a poor performance, when fed with the optimized solution.

(10) Strict Feasibility: The optimized solution satisfied all constraints – but other “solutions” have been preferred (e.g. designed manually by railway practitioners), although they violated some *less important constraints*.

(11) Reliability: The optimization software did not provide useful solutions on a regular basis (e.g. due to software bugs, or due to unreliable quality given that randomized elements have been deployed).

(12) Obsolescence: During the project duration, there have been new algorithmic findings which made the optimization methods in the project obsolete.

(13) Cost: The cost to make the optimization methods available in a productive context blast the cost which has been assumed in the cost-benefit-analysis that had been the basis to initiate the project.

(14) Attention: During the project duration, the “management attention” decreased, e.g.

because some protagonist within the railway company left the project.

(15) Others: These shall be specified by the respondents.

3 Realization of the Survey

For our survey, we used the online survey tool LamaPoll (2018). The survey had been designed anonymously, and it was only accessible with designated access codes. In total, we sent more than 98 access codes to both, mathematical optimization experts and managers or practitioners within railway companies. In addition, the authors filled four questionnaires about projects in which they were active. The geographical focus has been Europe (in particular Germany, the Netherlands, Switzerland, Denmark, Italy, France, Great Britain, Sweden), but we also asked experts from Northern America and China. The survey had been open from January 7th until January 21st 2019. In our inviting email, we were asking: “Hence, if ever in the past you had any project in which mathematical optimization had been intended to enter the practice of a railway company, but finally did not (fully) succeed, then we will be most thankful if you share with us your experience by answering the following questionnaire.”

We received 24 questionnaires, in which at least some of the problem specific features had been answered, including 22 questionnaires that had been finished, i.e. in which 100% of the mandatory questions had been answered. These include four questionnaires of the authors. Ten questionnaires had been filled by railway employees, and the other 14 by mathematical optimization experts. Moreover, since we had been interested in the personal experience of the protagonists, when we had been asked by two experts who were active in the very same project, we invited them to fill one questionnaire each.

We were also asking – optionally – for the projects’ names. Our intention was to possibly compare the answers of a railway manager on the one side, and an optimization expert on the other side, for the very same project. Indeed, in eight questionnaires the projects had been referred to with their names. But all projects had been different, so we are not able to perform such a comparison. Yet, this proves that the survey had not just been filled with ten questionnaires for the very same project.

Nevertheless, we are fully aware that N=24 is far away from letting us interpret the answers as being representative! Yet, we still consider the answers that we were able to collect as one step to provide possible explanations for the gap between the large total number of railway optimization projects, and the somehow limited number of both, true success stories from a fully practical point of view, and reports on project failures.

4 Results of the Survey

It had been our initial intention of the questions that we collected in Section 2 to be able to subdivide the answers on features that had been critical for the project’s success. Unfortunately, in view of just ten replies from railway managers, we do not consider it being appropriate to subdivide this small number of answers even further.

Let us shortly explain a somehow technical step that we did for our evaluation: In Section 4.4, we are going to consider correlations between framework properties of a project and the features that could have been critical for the overall practical success of an optimization project for a railway company. To this end, we translated the text answers that the participants were able to select into points:

For instance, for the question “Who has been mainly responsible for the R&D part?”, we defined a "scale" from three (University) via two (Research Institute) to one (Software Company). Similarly, for the feature "Input: What has been the structure of the input that was necessary to feed the mathematical optimization models?", we defined the following "monotone" scale:

(3) All data – except for optimization specific parameters – have been available in one existing IT system

(2) All data have been available in IT systems but had to be combined from more than just one system

(1) Some of the data that had been necessary to feed the mathematical optimization models have not been available in any existing IT system.

Now, we are ready to report on the answers given by the 24 participants. The scale of all the answers of Section 2.2 that we are reporting on in the sequel ranges from zero ("not relevant") to five ("decisive").

4.1 Most Decisive Features in the Eyes of Railway Managers

We start by providing the project features that railway managers and practitioners rated to be most critical for the practical success of an optimization project.

The Top 3 such features are:

- Attention: During the project duration, the "management attention" decreased, e.g. because some protagonist within the railway company left the project
- Validation: The railway company didn't allocate a sufficient amount of expert staff (time capacity) to validate in detail the results of the optimization methods during the entire project timeline
- Data: The available input data finally did not meet the quality that was necessary to be able to come up with high-quality optimization results (e.g. better than solutions that were designed manually)

In Figure 1, the problem-specific features of a project are ordered decreasingly according to the relevance that railway managers and practitioners associated with them on average why the developed methods did not enter practice on a regular basis. In addition, we display the range from the minimum value (light-gray, bottom) to the maximum value (light-gray, top), as well as the 25%-75% percentile (dark-gray).

We shortly comment on the Top 3 features. Regarding "Management attention", at least in business-oriented companies, let us have a closer look on projects that suggest a contribution to the company's benefit (e.g. by reduction of cost). Here, we believe that the management shall mainly be driven by economical goals, which typically can be expressed in terms of money. So, we believe that only *very* rarely, a decrease of management attention can be the *only* decisive feature if a project is terminated without entering practice on a regular basis. Rather, we fear that in most of the cases there might have been deviations from the initial profit estimate (higher cost for development/implementation, less savings for the application phase), too. For primarily service-oriented projects, if additional quality cannot be "translated" precisely into additional earnings, we are fully convinced, that a loss of management attention can be the initial cause for a project to be cancelled. Very much compatible to this consideration, let us shortly include the optimization experts' answers: In total, there have been 12 of 22 questionnaires, in which a "loss of management attention" had been rated (much) important, i.e. "4" or "5" – and *none* of these replies appeared in any of the 5 (of 22) projects, whose exclusive goal had been a reduction of cost.

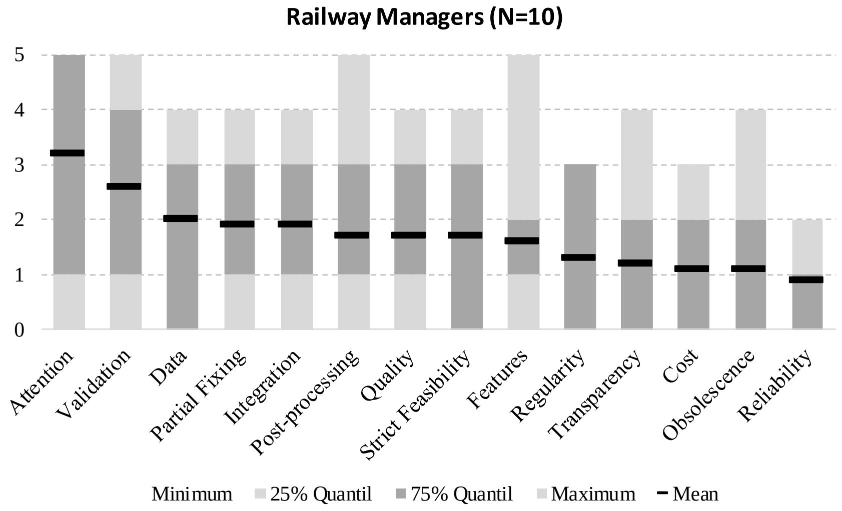

Figure 1: Relevance that railway managers and practitioners give to the problem-specific features of a project why it did not enter practice on a regular basis. For instance, "Data" had been given an average relevance of 2.0, a maximum of 4, and the 25%-75% Percentile ranges from 0 to 3

To be honest, we have been surprised in a positive way, that also railway managers seem to be aware that a shortage of expert capacity for validation – and thus, in the sequel, for the improvement of software prototypes – can indeed be a decisive feature for the unsuccessful end of a project. Nevertheless, optimization experts associate with it an even larger relevance, see also Sections 4.2 and 4.3.

In Section 4.4, we will see that the fact that the availability and/or quality of input data has been rated with 3 or 4 by 50% of the railway managers in particular correlates with a general feature of the project, namely the IT system environment on the input data side. However, recall that when discussing correlation, we are still aware that N=10 and N=24 are not suited to guarantee any true statistical significance.

4.2 Most Decisive Features in the Eyes of Optimization Experts

Now, in Figure 2 let us turn to the perspective of the mathematical optimization experts. Much like the railway managers and practitioners, they rated the shortage of expert capacity for validation being relevant – but with a by far more striking average of 3.64 out of 5.

Among the Top 5, here we also get what we called "strict feasibility", for short, i.e., the fact that in the end practitioners might have made use of the possibility to "relax" some of the constraints that have been imposed to the optimization algorithms, still considering their manually designed "solution" to be "practically feasible".

In addition, mathematical optimization experts consider the cost for making the optimization methods available in a productive context relevant, if they exceed the initial cost-benefit-analysis of the project (2.50).

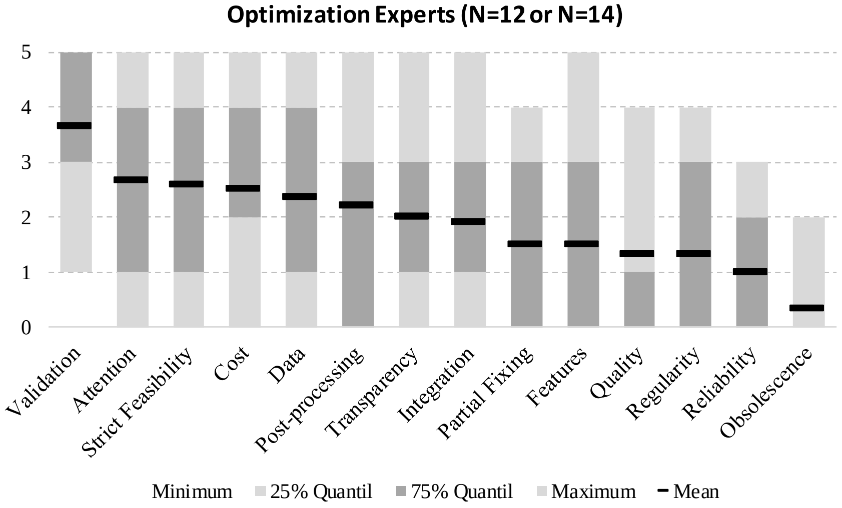

Figure 2: Relevance that mathematical optimization experts give to the problem-specific features of a project why it did not enter practice on a regular basis.

4.3 Features with Largest Deviations Between Managers & Optimization Experts

Even more interesting insights might arise from comparing the relevance that the railway managers and practitioners assigned with the one that the mathematical optimization experts did assign. In Figure 3 we subtract the mean of the latter from the mean of the former.

For any of the differences that are displayed in the chart, keep in mind that we have to assume that there is *no* project, for which we got the answers from both sides, i.e., railway managers and optimization experts. Hence, the primary reason for any differences between the two perspectives could still simply lie in a different nature of the projects. Nevertheless, assuming that the major source for the difference could indeed lie in the role of the protagonists, we propose the difference values as a kind of indication.

At first sight, one could observe that from the perspective of the railway managers, obsolescence of the algorithm appears to be much more relevant for a project not to attain its full goals, compared to the understanding of the mathematical optimization experts with respect to *their* methods. But recall from the previous figures that the values for the feature “obsolescence” are 1.1 for the railway managers and practitioners, but only 0.3 for the mathematical optimization experts, which yields the value 1.1–0.3 = 0.8. In particular, both partners did only observe a (very) small relevance in “obsolescence” of the methods.

At the other end of the scale, it had been the impression of the mathematical optimization experts that the full integration of their methods into the software landscape of the company turned out to be too costly in the end, and thus become a “show-stopper”. This is reflected by a value of 2.5. Interesting enough, this is not confirmed by the railway managers and practitioners, who rate this feature only 1.1, which thus provides a difference of 1.1–2.5 = –1.4.

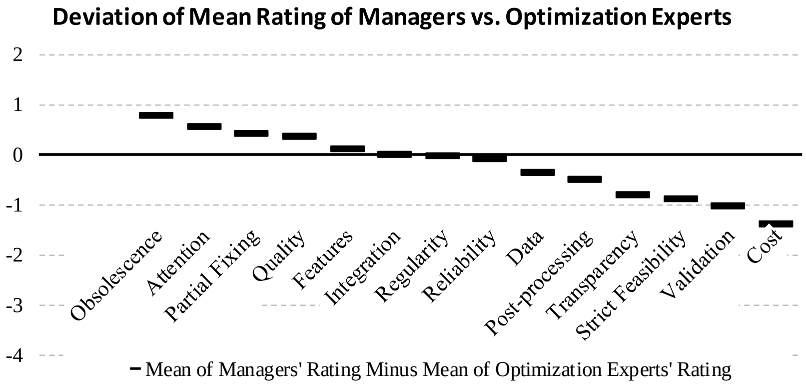

Figure 3: Difference between the mean rating assigned by the railway managers and practitioners and the mathematical optimization experts

Although the feature “Validation” (expert capacity to (in-) validate intermediate results) belongs to the Top 2 features of both, railway managers and optimization experts, there is one of the largest *gaps* between the intensity that they assigned to this feature: 2.6 by the railway managers, but even 3.6 by the optimization experts, hereby marking the top score of the entire survey. This provides a difference of 2.6–3.6 = –1.0.

Moreover, we find another result interesting. Consider the “moving target” question (f), which we put in the “general framework” section of our questionnaire. Only 20% of the railway managers said there has been a “moving target” within their projects – while as many as 57% of the mathematical optimization experts report this as their impression! We suppose that this could be due to different understandings regarding the degree of specification at the very beginning of a project: Maybe, railway managers sometimes cannot (or do not want to?) specify any requirement in most detail when launching a project. Then, later, when they “add” some piece of specification, the mathematical optimization experts could experience such a late specification already as a significant modification of the project goal, or “moving target”.

4.4 Selected Correlations Between General Framework and Specific Features

Finally, although we are fully aware that 10+14 filled questionnaires unfortunately cannot be representative for all projects and project members, we still perform some correlation test and invite the reader to interpret it as a slight indication.

To this end, we computed the correlation between each pair of feature of the general administrative framework of the projects (see Section 2), and of the problem-specific properties of the projects (see Section). Among the roughly 250 possible combinations, one can detect three where the absolute value of the correlation is larger than 0.5, and thus could tend to be “significant” (which it is not, due to our relatively small sample size).

- 0.65
  The more academic the partner who has been mainly responsible for the R&D part (3 = university, 2 = research institution, 1 = software company)…
  … the more severe the lack of railway expert capacity for validation for the situation when a mathematical optimization project does not meet its full

goals.

- –0.55
  The better the structure of the input data (3 = all data available in *one* existing IT-system, 2 = all data available in some IT-system, but these must be combined, 1 = some of the data necessary for the mathematical optimization models have not been available in any existing IT system)…
  … the less likely the project's results did not get used in practice on a regular basis because the available input data finally did not meet the quality that was necessary to be able to come up with high-quality optimization results.
  Of course, this relationship sounds absolutely reasonable. So we primarily interpret the observed correlation as a kind of cross-check question to evaluate the consistency of the answers, rather than some new insight.
- 0.54
  The fact that the project goal has been modified significantly during the project ("moving target") correlates positively with…
  … the priority that had been assigned to the fact that in the end solutions that had been designed manually by railway practitioners made their way into practice, although they violated some (less important) constraints which the optimization software still had to respect ("strict feasibility").
  A similar positive correlation (0.48) can be observed between the "moving target" property in the general framework, and the project-specific feature "regularity" (The optimization results didn't show a certain "regularity pattern", which in the end had been expected by the railway practitioners, although not communicated as a key feature at the project kick-off).
  Also here, we consider these two correlations very much reasonable: A moving target and either of "regularity requirements" (not communicated at the very beginning) and "strict feasibility" (relaxations at the very end) can be regarded as two sides of the same coin. This makes us believe in the quality of the answers that we received, despite their small number.

To summarize, we were able to statistically observe some correlations between features that we were asking in the context of "general framework" of a project, and problem-specific reasons why a project finally did not enter practice on a regular basis. While the second and the third one that we are reporting on are rather confirming somehow trivial assumptions, the first one might constitute a "lesson learnt": In particular, when the R&D part is contributed by a university, it is even more critical for the actual success of the entire project that the railway company allocates a sufficient amount of capacity of practical experts in order to evaluate intermediate results.

4.5 Further Comments by the Experts

The last – optional – question of our survey has been: "Have there been other features for the project finally not to meet its full goals?", i.e. those, which our questionnaire did *not* include already (see Section 2.2). In the sequel, we report some answers that we received for this question:

- Complexity of Control
  In a sense symmetric to a lack of transparency of the solution, it can also be negative for an optimization tool, if it leaves too many control parameters to the end user, where the effects (and interactions) of the parameters could not be anticipated adequately.

- Employee Participation
  In a project that touched the working times of staff, the best solutions that the optimization tool was able to deliver have not been accepted by the unions, and thus decreased the benefit of the project, and/or delayed its implementation significantly.
- Management Implementation
  The change required towards an automated optimization method was significant: therefore, a relevant contribution from the management would have been required to make it used in real practice.
- Managerial Consistency
  The gap between the management expectation to substantial benefits on the one hand vs. the very detailed “parameter battle” with the experts on the other side could not be closed. The correct parameters have an extremely high impact on the optimization result. Therefore, a lot of time of the railway experts is needed (see also “Validation”). This could not be communicated to the management.
- Organizational Changes
  Suboptimization within organizations and organizational units meant the global optimum provided required large organizational changes to be implemented in practice.
- Performance
  The runtime of the optimization was much higher than expected. The optimization approach used wasn’t suitable for the size of the problems as it is relevant in practice. If the scientist is able to deliver a high quality solution after a computation time of 48h, then it is only of limited use for a practitioner, if he requires the results in a „live“ context.
- Rolling Horizon
  If a shift plan had to be designed for some general week, it should of course “glue well” between Sunday 23:59 and Monday 0:00, without leaving an expensive transition back to the initial state outside the objective function.

5 Conclusions

Even though the number of replies that we received stayed rather small, we feel able to provide some suggestions for the future design and management of operational research R&D projects for railway companies. Recall that here we are not referring to projects, in which just some study for the potential of some new algorithmic ideas is to be conducted. Rather, we are considering projects that have the goal, that at the end the optimization methods will be used in practice on a regular basis.

- The by far most reported reason why in the past the results of optimization projects for railway projects did not enter practice on a regular basis, is a lack of *expert capacity within the railway companies for the validation of intermediate results*. An appropriate amount of their capacity must be planned from the very beginning of the project, and then guaranteed throughout the lifetime of the project.
- This point has been rated even more important, if the R&D part in the project has been developed by a university partner – presumably, software companies

fixed the required capacity allocation already in their contracts? In any case, we encourage in particular university partners to do so for future projects.

- The *availability, consistency, and quality of input data* can of course be decisive for the success or failure of any project. Hence, we recommend in particular to the railway companies to let their R&D partner evaluate the quality of the available input data in detail prior to launching the actual project for the development of algorithms. If there were some significant deficiencies detected, then it could make sense to postpone the optimization project until the input required for it is available.
- Regarding *management attention*, let us only consider quality-oriented projects, where the contribution to the benefit of the company cannot be expressed explicitly (in terms of money). We agree that management attention risks to be volatile in particular in this case. Here, we can only recommend to the companies only to initiate such projects, of which they can be sure that their (strategic?) quality goals will *not* change during the timeline of the project.
- Finally, let us recall the "*moving target*" property of a project, which we observed to be much more present in the eyes of mathematical optimization experts. To prevent a project to fail due to this feature, we recommend to the railway managers to put very much emphasis on a detailed description of the requirements for the optimization tool, and prevent any deliberate "lazy specification".

6 Acknowledgements

We are most thankful to the 24 respondents of our survey, to another half a dozen experts who answered our open "other features" question whose answers are reported in Section 4.5, and to Verena Appeldorn (DB Cargo AG), whose feedback to an earlier initiative laid the ground for this survey.

References

Borndörfer, R. et al., 2017. "Recent success stories on integrated optimization of railway systems", *Transportation Research Part C: Emerging Technologies*, vol. 74, pp. 196 – 211. https://doi.org/10.1016/j.trc.2016.11.015

Borndörfer, R. et al. (eds.), 2018. *Handbook of Optimization in the Railway Industry*, ISBN 978-3-319-72152-1, Springer International Publishing.

Cacchiani, V. and Toth, P., 2012. "Nominal and Robust Train Timetabling Problems", *European Journal of Operational Research*, vol. 219, pp. 727 – 737. https://doi.org/10.1016/j.ejor.2011.11.003

Caimi, G. et. al., 2017. "Models for railway timetable optimization: Applicability and applications in practice", *Journal of Rail Transport Planning & Management*, vol. 6, pp. 285 – 312. https://doi.org/10.1016/j.jrtpm.2016.11.002

Gorman, M., 2016. "From Magnum Opus to Mea Culpa: A Cautionary Tale of Lessons Learned from a Failed Decision Support System", INFORMS Journal on Applied Analytics, vol. 46m pp. 185-195. https://doi.org/10.1287/inte.2015.0818

Harrod, S. and Gorman, M.F., 2011. “Operations Research for Freight Train Routing and Scheduling”, In: Cochran, J.J. et al. (eds.), *Wiley Encyclopedia of Operations Research and Management Science,* ISBN 978-0-470-40063-0, John Wiley & Sons, Inc.

Kohl, N., 2003, “Solving the World’s Largest Crew Scheduling Problem”, *orbit Xtra*, pp. 8 – 12.

Kroon, L. et al., 2009. “The New Dutch Timetable: The OR Revolution”, *Interfaces,* vol. 39, pp. 6 – 17. https://doi.org/10.1287/inte.1080.0409

LamaPoll, 2018. An online survey tool by Lamano GmbH & Co. KG, Berlin, Germany. https://www.lamapoll.de/Support/Impressum

Liebchen, C., 2008. “The First Optimized Railway Timetable in Practice”, *Transportation Science*, vol. 42, pp. 420 – 435. https://doi.org/10.1287/trsc.1080.0240

Schülldorf, H., 2018. “Der weite Weg der Optimierung zum produktiven Einsatz“, presentation at the *Sitzung GOR - Arbeitsgruppe Logistik und Verkehr*, May 18th, 2018, Karlsruhe, Germany, in German.