2017:00898 - Unrestricted

Report

InfraPlan

A tool for socio-economically optimal railway infrastructure investments under uncertainty

Author(s)
Truls Flatberg
Marte Fodstad
Michal Kaut
Arnt-Gunnar Lium
and Adrian Werner

Photo by David Gubler / CC BY-SA

SINTEF Technology and Society
Applied Economics
December 21, 2017

SINTEF

SINTEF Teknologi og samfunn
SINTEF Technology and Society

Address:
Postboks 4760 Sluppen
7465 Trondheim
NORWAY

www.sintef.no

Enterprise Number: NO 948007029 MVA

KEYWORDS:
Rail transport
Cost-benefit analysis
Investment planning
Optimization
Stochastic programming

Report

InfraPlan

A tool for socio-economically optimal railway infrastructure investments under uncertainty

VERSION
-

DATE
December 21, 2017

AUTHOR(S)
Truls Flatberg
Marte Fodstad
Michal Kaut
Arnt-Gunnar Lium
and Adrian Werner

CLIENT(S)
Jernbanedirektoratet

CLIENT'S REFERENCE

PROJECT

NUMBER OF PAGES AND ATTACHMENTS
14

ABSTRACT
Cost-benefit analysis is the de facto standard in Europe for studying investments in railway infrastructure. This report describes a model extending this approach to aid real-life decision making under budget and technology uncertainty. The model, InfraPlan, has been developed for—and in cooperation with—the Norwegian railway infrastructure manager, Jernbanedirektoratet (Norwegian Railway Directorate), and builds upon their current methodology for cost-benefit analyses. We extend traditional cost-benefit analysis into an investment-project portfolio model based on multi-stage stochastic programming. We enrich the traditional approach with ideas from portfolio analysis, network design, line planning and transportation demand modeling, overcoming the shortcomings of planning each aspect separately. A thorough appraisal of investment alternatives is important to ensure both fair comparison between projects and efficient use of public money. In this report, we present a model that can consider a portfolio of investment alternatives. We also demonstrate how it can be used to increase the benefit from railway infrastructure investments, finding an optimal selection of options that improves cost efficiency and, ultimately, can reduce risk.

PREPARED BY
Truls Flatberg

SIGNATURE
Truls Flatberg

CHECKED BY
Frode Rømo

SIGNATURE
Frode Rømo

APPROVED BY
Vibeke Nørstebø

SIGNATURE
Vibeke S Nørstebø

REPORT NUMBER
2017:00898

ISBN
978-82-14-06603-6

CLASSIFICATION
Unrestricted

CLASSIFICATION THIS PAGE
Unrestricted

Contents

1 Introduction

We consider a railway system with separation between the, usually public, infrastructure owner and manager (IM) and train operating companies using the infrastructure, such as is common in Europe. Focusing on infrastructure investment analysis, we take the IM's perspective. Investments in railway infrastructure involve high costs, long planning and construction times and a typical lifetime of several decades—potentially up to several centuries for investments such as tunnels.

An investment in one part of the network may affect not only infrastructure in close proximity, but also distant or apparently not connected parts. Consequently, it may have an impact on how the network is used in other areas and, hence, on the utility or return on investment of these infrastructure parts as well. This means that, when performing an investment analysis, one needs to consider the infrastructure as a whole rather than study investment options on infrastructure parts in isolation.

Moreover, there is significant uncertainty related to the demand for the infrastructure and its usage over its long lifetime. Uncertainties about the quality of the infrastructure, changes in capacity due to technological development (e.g., improved dispatch automation), cost of construction and use or about future budgets complicate an investment analysis further. Current planning typically approaches the problem in a deterministic fashion, resulting in investment suggestions that are suboptimal or too optimistic in terms of their ability to deal with unforeseen development directions.

Public infrastructure investment plans are frequently based on cost-benefit analysis (CBA) as it is known from economic theory, see for example Boardman et al. (2014). However, such economic analysis traditionally focuses on evaluating effects of *individual* investments rather than coordinating several investment opportunities—a limitation pointed out already by Mackie and Preston (1998). A parallel tradition in the operations-research community addresses this through network-design models (Magnanti and Wong, 1984) or their stochastic versions (Bai et al., 2014). Such models can be used to determine where to invest in new infrastructure, in which type, when and with which capacity, to satisfy some demand at the lowest cost possible. Lin et al. (2012) apply network design models to optimize freight train connections in a large-scale rail system.

In the railway sector, planning has traditionally been done in a deterministic and sequential fashion, where analyses and decisions at one level form the basis for and constrain planning at subsequent levels: Results from CBA or network design models establish the basis for line planning, which determines the lines or services and their frequency of operation. Next, rolling stock plans are developed to determine which services should be operated by which type of rolling stock and how many units are required. Finally, tactical and operational planning is concerned with timetabling, rolling-stock rostering, crew planning and traffic management.

As an entity with a public mandate, the IM should be concerned not only with the immediate economics of infrastructure investment plans, but also with their wider implications in a socio-economic context. This entails considerations of how the infrastructure can be used best, taking into account its respective state. In other words, infrastructure planning should be tightly connected with planning at all other levels. There exists a significant body of literature on how to address the various planning levels. However, most of the work considers only one specific level. Frameworks such as the one suggested by Bussieck et al. (1997) iterate multiple times across levels to improve plans, but they are usually not applied in practice due to the high degree of complexity and resources required for each step.

Mackie and Preston (1998) highlighted several challenges related to transport-project appraisal. While we do not address all, our approach overcomes the biggest limitations: 1) planning is done project by project, for a limited geographical area or separately for each level, 2) with CBA, only a limited number of alternatives or combinations is evaluated, and 3) uncertainty is handled ex-post, for example using simulations, instead of being taken into account during the project-selection process.

The supply-chain literature, such as Thomas and Griffin (1996), provides numerous references for how multiple echelons of supply chains can be integrated to enable better solutions. From finance (for example, Zenios (2007)), we know that a portfolio of investments can help to reduce risk or to increase expected return on investments. The operations-research community has made methodical contributions to optimizing investment decisions in time and space and for a large number of options. Surprisingly little work has been done on how to combine the rich framework of economic analysis with tools from the field of optimization under uncertainty such as stochastic programming (Birge and Louveaux, 2011) or stochastic dynamic programming (Bertsekas, 1987).

Bridging the gap between economics and stochastic optimization, the InfraPlan model presented in this report is an economic framework for real-life infrastructure-investment decisions, formulated as a multi-stage stochastic-programming model. It can take into account both budget uncertainty and uncertainty related to the cost and performance of new technology, allowing for learning effects from investing in new technology. In order to assess effects of the investment decisions, the model incorporates elements from network design, line planning and train-schedule generation. Demand response to the new infrastructure and the offered services is an integral part, which is addressed using demand elasticities.

Stochastic-programming approaches to find optimal investment portfolios under uncertainty, taking into account operational aspects and sector specific detail, have been developed for natural gas transport infrastructure (Hellemo et al., 2013; Fodstad et al., 2016), liquid natural gas transport and storage (Werner et al., 2014), maritime fleet size and mix problems (Gundegjerde et al., 2015) and other sectors. The framework presented here is, to the best of our knowledge, the first example of an application to railway networks.

This report is a result from two projects funded by the Norwegian infrastructure manager Jernbaneverket (now Jernbanedirektoratet). In the following section, we discuss some background aspects motivating our approach. Section 3 gives an overview of the InfraPlan model and how it addresses the specifics of railway infrastructure investments. We demonstrate selected aspects of our model in a case study in Section 4, illustrating typical situations where InfraPlan will be beneficial, and conclude in Section 5.

2 Using optimization to coordinate investment decisions

In economics, significant contributions have been made to develop methods improving public transportation infrastructure. While such improvements may not be profitable by itself, other effects may still make them socio-economically beneficial. An important aspect in this regard is potential travel-time savings due to, e.g., investments leading to increased train speed and frequency. Another effect is that railway infrastructure development can contribute to shifting traffic from other transport modes, which has environmental benefits and may reduce costs associated with road and other transport infrastructure (Gorman, 2008). A typical cost-benefit approach includes also aspects such as the value of fewer deaths and injuries in traffic, reduced congestion, decreased noise and pollution, as well as non-valued costs and benefits such as scenery. In the case of European railway transportation, each country has its own guidelines for performing CBA (Odgaard et al., 2006) and one may arrive at different conclusions due to differences between the guidelines (Olsson et al., 2012).

In practice, an IM often does not analyze a single investment option in isolation. Rather, various options for several railway lines are investigated with a planning horizon of multiple decades, on a more or less continuous basis. These investments can be of various kinds, such as construction or upgrades of tracks (e.g., electrification), stations, terminals or parking areas, but also dispatch and control systems. Maintenance of the existing infrastructure comes in addition. Having to consider such diverse investments complicates the analysis significantly, as they may have different, even opposed or overlapping, effects.

As an example, consider the choice between upgrading parts of a line from single to double track and investing in an advanced train dispatching system. Both investments can be expected to increase capacity and improve flexibility in an operational setting. However, the former option has a long expected lifetime and will require maintenance for as long as it is in use, while the latter has a shorter expected lifetime but is likely to reduce the train operators' energy costs. It also comes at a greater risk as the technology is less mature. Moreover, the alternatives can be combined and one has to decide on their timing. For example, delaying an investment in the dispatch system may reduce risks associated with the technology's immaturity.

For realistic analyses, complexity grows fast with the number of possible combinations of investments, in particular when taking into account timing aspects or dependencies between options. This holds even more when considering network effects (changing one line affects other lines as well) or budget constraints, which are often disregarded in CBA. Also, for each considered investment strategy, one should traverse further planning levels as discussed in the introduction, find good responses to the given strategy and then take an overall assessment of the strategy's quality. That renders traditional CBA approaches of evaluating all alternatives separately impractical, even prohibitive, as one can analyze only a limited number of investment strategies and select the best one(s) among those. This risks overlooking the best solutions.

An optimization-based approach, on the other hand, allows a more structured analysis by formulating all effects, relations and limitations, also temporal ones, as well as the goals in one model. This allows the inclusion of several planning levels, timing aspects or complex dependencies between decisions. Moreover, by employing, e.g., stochastic-optimization concepts, uncertainty can be taken into account, which would have been ignored in a pure CBA approach. Systematically exploring potential investment strategies and related decisions, a suitable solver is then able to find the guaranteed best solution satisfying all conditions. The solution time depends on problem size, structure and complexity, but for the InfraPlan model discussed here, realistic problems are typically solved within minutes or up to a few hours. Consequently, the efficiency of an optimization-based approach allows both more in-depth and more comprehensive investment analyses than current practice.

However, one should bear in mind that, while optimality can be guaranteed even for fairly large problems, the approach can take into account only effects which can be quantified in a sensible way. Like for traditional CBA approaches, non-valued aspects should, hence, be assessed manually as part of the general work process.

3 The InfraPlan model

In this section, we present the InfraPlan optimization model. We focus on the most important features since discussing the whole model, with all its details, is beyond the scope of this report.

In the context of the model, all considered investment opportunities are referred to as *projects*. There may be various types of projects: construction of new tracks, stations and terminals or modification of existing ones, upgrade of existing tracks to increase the maximum allowed speed, upgrade of stations and terminals to increase their capacity, electrification of lines, and so on. Each project is defined by start time window, completion time, location, and the impact it has on the railway infrastructure. The model then decides whether the project should be started and, if so, when. In addition, we incorporate simplified line / service planning by allowing the model to decide on the train frequency of the involved lines.

Similar to most CBAs, and unlike standard optimization approaches, the model assumes a reference alternative. There is no unambiguous way of determining such a reference. Typically, it represents an investment plan that consists of no or few investments or just the investments required to maintain the status quo. The model then finds a portfolio of projects that maximizes expected net benefit, relative to the reference, over the given planning horizon.

3.1 The objective function

The goal of the model is to maximize the expected net benefit of the investments over the planning horizon. Both benefits and costs are calculated relative to the given reference alternative and are expressed in monetary terms with their expected net present value as common basis. Thus, the objective is to find the socio-economically optimal solution considering the whole project portfolio.

The cost side consists primarily of the investment costs for started projects. The benefit calculations are more complex and can be divided into four categories: benefits for train users, train operators, the IM, and third parties. Benefits for third parties are related to accidents and environmental effects; increasing train traffic will lead to higher expected accident costs and emissions, both locally (e.g., noise) and globally (e.g., CO_2 from diesel trains). Benefits for the IM include changes in costs for operating and maintaining the infrastructure. Benefits for train operators are normally not considered as any changes in their costs should be compensated by corresponding changes in public procurements or contracts with their customers. These benefits are estimated only based on changes in the number of trains.

To calculate benefits for train users, both passengers and freight, we also have to take into consideration the traffic between origin–destination (O–D) pairs. The benefits for passengers are calculated as changes in their generalized costs, incorporating fare and the passengers' valuation of waiting and travel time. Waiting time depends on the frequency of the services available for each passenger and can, for example, be reduced by projects that increase track capacity. Similarly, travel time can be reduced by projects that allow higher train speed. To find the total benefits, the model keeps track of the number of passengers on each possible route between each O–D pair and of the changes in generalized cost for each route. For freight customers, the benefits are primarily related to changes in freight costs. These costs are calculated by finding the equilibrium cost where offered capacity equals demand.

3.2 Infrastructure usage and capacity

In our model, railway infrastructure is represented by a graph with nodes and arcs. The nodes represent stations, shunting yards and freight terminals while arcs, corresponding to tracks between the stations and various facilities, connect the nodes.

In general, the capacity of the infrastructure depends on many factors such as number of tracks, maximum allowed speed, station lengths, number of calls at stations or terminals, mix of train types (acceleration profiles, maximum speed, etc.) or direction of the lines. It is, therefore, often determined by way of simulations or the like. In InfraPlan, capacity utilization of an arc is expressed through the capacity used by each train and the number of trains traversing that arc, constrained by the maximum available capacity on the arc.

Capacity can be reduced through wear and tear which may necessitate speed reductions to maintain safe operations. However, it can also be increased through investments establishing new tracks or increasing speed on existing ones.

Figure 1 shows a real-life example involving the Dovre and the Gjøvik Line, just north of Oslo. Solid black lines represent existing tracks and connect nodes indicating selected stations. Capacity on the Dovre Line around Oslo (*osl*) is insufficient, leading to congestion and operational challenges, not least since the tracks have to be shared with long-distance and goods train services. There are plans to continue upgrading the Dovre Line to double tracks until Moelv (*moe*), and potentially until Lillehammer (*llh*). The construction of a connector between Gjøvik (*gjø*) and Moelv, indicated by the dashed black line, is also considered. This will free up capacity on the Dovre Line between Oslo and Moelv stations as some services can be routed via the Gjøvik Line and the connector.

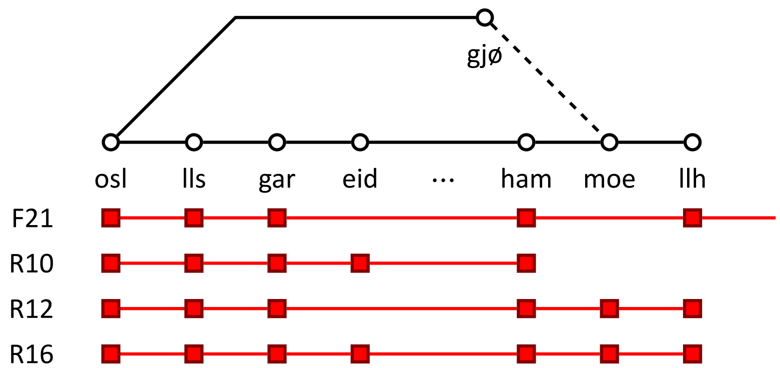


Figure 1: Existing infrastructure, potential connector and examples of passenger train services on the Dovre and the Gjøvik Line.

Key questions in this respect are: 1) Should one invest in additional capacity? 2) Which project(s) should be chosen and in which sequence? 3) If investing, when should construction work start and be completed? 4) How may carrying out one such project affect decisions on the other plans? For example, would decisions on upgrading the Dovre Line be different if the connector had already been constructed?

The tracks are used by various passenger and freight train services, some of which are indicated in red in

Figure 1. Regional passenger services call only at the largest stations, while local trains stop at every station. Also terminal stations may vary for the different services. Regional services are operated by faster train units with large capacity in terms of passenger kilometres per hour, while the slower local services have lower capacity. Increasing the number of tracks may free up capacity also by changing the train mix, for example, by routing some regional and freight services via the new infrastructure or allowing for new timetables to be introduced. As capacity utilization does not only depend on the number of trains relative to the line capacity but also on their combination, the questions concern not only investing, but also how to best use the infrastructure whether an investment is done or not. The InfraPlan model accommodates this aspect through a simplified assessment of the investment options' impact on service frequency—and thus capacity utilization and satisfied transportation demand.

3.3 Demand modelling

As demand for passenger and freight transportation varies throughout the day, service frequency may vary accordingly. For example, local passenger trains stopping at all stations may have a high frequency in morning and afternoon rush, while freight trains stopping only at terminals operate with high frequency at night and low frequency during the day.

The demand is described by time-dependent O–D matrices, allowing to express different demand levels at peak and non-peak hours (which may be different times for passengers and freight). Passengers and freight can be transported on one or more lines as long as these connect the appropriate O–D pair. Further, for passenger transportation we distinguish several types of travellers, such as commuters and business travellers, in order to use a Ramsey–Boiteux pricing approach.

The O–D matrices are based on the existing infrastructure and the level of service in the reference alternative. Improving or degrading the level of service affects the demand. For passenger transport, we estimate this change by way of generalized costs and the demand in the reference alternative in combination with the price elasticity of demand.

The generalized costs include, in addition to ticket price, components for waiting time, travel time, delay time and crowding. These aspects may be valued differently by the different traveller types. Obviously, demand will increase when generalized travel costs decrease. As long as capacity on board the trains is sufficient, all demand can be satisfied. However, if the demand increase leads to overcrowded trains, generalized travel costs will increase and potential passengers may choose to not travel by train. We model the cost of crowding by way of a piecewise linear function taking into account seat and total capacity of the trains as illustrated in Figure 2.

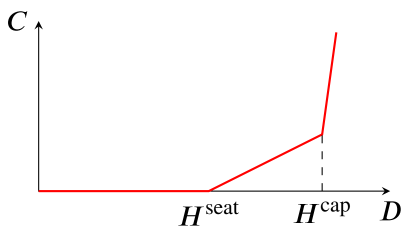
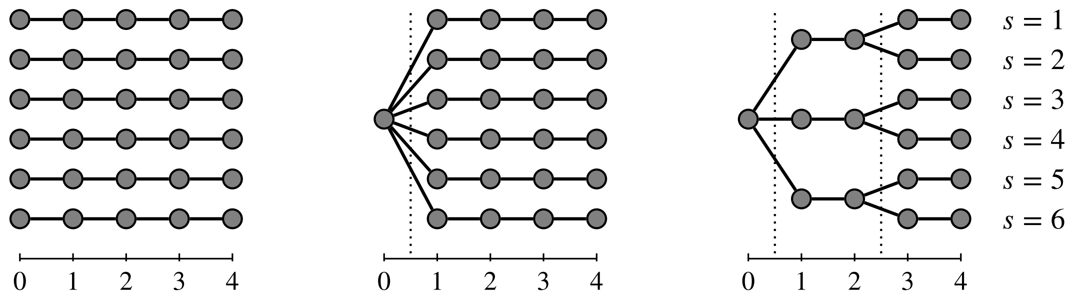


Figure 2: Cost of crowding modeled as a piecewise linear function with breakpoints at the seat capacity H^{seat} and at the maximal capacity H^{cap}.

For freight transport, the amount transported between an O–D pair is limited by the maximal demand given by the O–D matrix and the offered transport capacity in terms of daily departures and train capacity. If capacity is the limiting factor, it is assumed that the train operating companies will increase prices, leading to a decrease in utility for freight customers. This, in turn, will result in a decrease in transportation demand – until demand equals offered capacity with a corresponding surplus for the company.



(a) A set of scenarios **(b)** A two-stage tree **(c)** A multi-stage tree

Figure 3: Three scenario-tree structures, each with 5 time points ($t \in \{0, \dots, 4\}$) and 6 scenarios ($s \in \{1, \dots, 6\}$). Nodes represent decision points while arcs represent realizations of uncertain data. The dotted lines show borders between stochastic 'stages', that is, parts of the model having the same information.

3.4 Quality and degradation

The state of the railway infrastructure is modelled by one or more quality dimensions, which can be affected by usage and investments. The quality impacts the cost of upkeep of the infrastructure, how trains can be operated safely, on time, with high regularity (i.e. few cancellations) or whether specific train types such as long freight trains or electric trains can be used.

Quality can decrease due to wear and tear or just decline over time. In the former case, it is related to the number of train movements and/or the type of trains. A heavily loaded freight train will wear down tracks and ballast much more than a light local passenger train. For other components such as electronics, the quality is usually not dependent on how much they are used, but rather on the duration of their usage: After some time, it is reasonable to expect that the probability for failure or the cost of upkeep will increase significantly.

Reliability engineering frequently describes hazard functions of railway infrastructure assets by a 'bathtub curve' (e.g., Klutke et al., 2003, and references therein). However, as the InfraPlan model has been developed for a long planning horizon, we disregard initial-phase failures and assume increasing and convex failure rates, reflected by concave quality-degradation functions. Further, due to lack of data as well as for computational tractability, we do not take uncertainty in failure rates into account.

Investments can bring the quality of an infrastructure element back to its original state, or to a higher level. For example, major maintenance of an old single-track line can restore its original state. Replacing it with a new double-track line can set capacity and regularity to a higher level.

The quality term is also used to describe aspects related to infrastructure that are binary in nature, such as electrifying tracks or having passing loops for long freight trains. In contrast to other quality dimensions, these typically do not decline. Moreover, the benefits they enable can only be obtained if the dimension applies to the whole of a line, that is, if all required investments have been carried out. For example, in order to operate a line with an electric train, all track segments used by the line must be electrified.

3.5 Uncertainty

To take uncertainty into account, we follow standard practice in stochastic programming and represent it in terms of a *scenario tree*. This means that we have a set of scenarios $s \in \{1, \dots, S\}$ and all model entities that would be indexed on time t in a deterministic model have to be indexed on the time-scenario couple (t, s) instead, reflecting potentially different values. This, however, would not be sufficient, as it would give us a set of S disconnected paths, as illustrated in Fig. 3a. In fact, this would mean solving S independent *deterministic* problems—akin to sensitivity analysis—and would result in S possibly different results.

To create a stochastic model, we require that the model has only one answer for $t = 0$, representing 'now'. In other words, we need the scenario-tree structure from Fig. 3b. In the model, we then have to ensure that all the decision variables at $t = 0$ have the same value in all scenarios. This can be achieved with extra constraints, usually called *non-anticipativity* (or *implementability*) constraints.

While the scenario tree from Fig. 3b represents a stochastic problem, it is only a so-called 'two-stage' problem, where by stages we mean periods with new information: in the first stage, covering $t = 0$, we do not know which scenario will happen, while in the second stage, covering the rest of the time horizon, we know it. In other words, from $t = 1$ onwards, there is no more uncertainty left (no more branching in the tree).

To get a more granular revealing of uncertainty, we need a 'multi-stage' tree, such as the one depicted in Fig. 3c. This tree has tree stages, the first at $t = 0$, the second at $t \in \{1, 2\}$, and the final stage covering $t \in \{3, 4\}$. In other words, we have obtained new information about uncertain values before $t = 1$ and $t = 3$.

If the uncertainty is in reality revealed continuously, having more stages should result in a better approximation of the learning process. On the other hand, the size of the tree grows exponentially with the number of stages, so one has to strike a balance between the quality of the approximation and the solution time of the resulting problem.

The InfraPlan model has been formulated for a general multi-stage scenario tree with scenarios $s \in \mathcal{S}$ with probabilities p^s, $\sum_{s \in \mathcal{S}} p^s = 1$. It should be pointed out that this approach captures only exogenous uncertainty, i.e., uncertainty that is not influenced by the model decisions. Modelling uncertainty that depends on decisions is significantly more difficult, as we show in Section 3.7.

3.6 Learning effects and uncertainty when introducing new technologies

The concept of learning curves, well-known in economics, addresses the effect experience has on efficiency (eg., Arrow, 1962). While it is often used to describe how production costs decrease with increasing experience, similar effects can be observed with the maturing of new technology where costs decrease and performance improves. Another effect related to maturing processes concerns uncertainty about a technology's performance, which reduces significantly with its usage. Hence, subsequent installations of the same type of technology can often be done at a lower risk than the initial one.

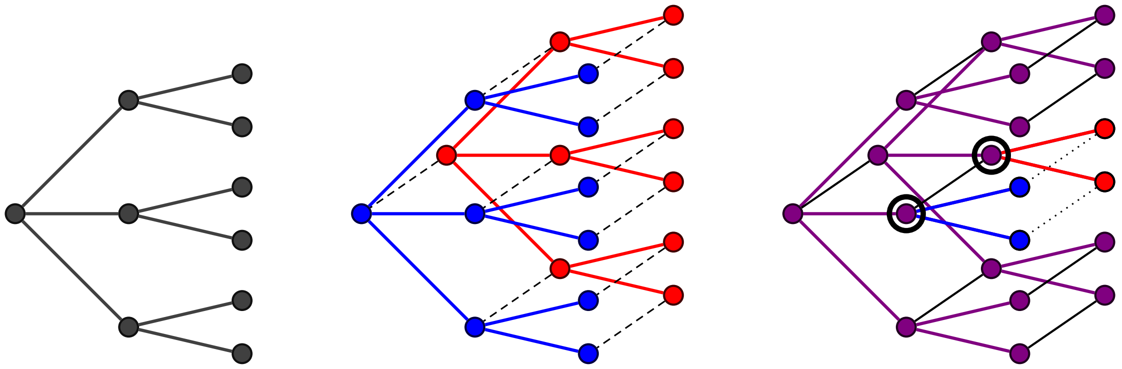
Emerging technologies in the railway sector are, for example, new types of trains, automated dispatching and driver advisory systems, and signalling systems (ERTMS). While they typically have been tested for some time in limited areas before being rolled out more generally, it may not be clear how they will perform in the long run (several decades) and under different conditions, nor how they will affect the overall performance of large-scale real-life railway systems.

On a larger network, there are two obvious approaches to introducing a new technology. It may be implemented first where it will have the largest positive impact, or we may start where any potential problems will have least negative impact. The former has higher expected values, but also significant risk to operational performance. The latter variant allows for gathering experience with the technology in a low-risk area and therefore reduces uncertainty in the consecutive investments in other parts of the network. Since we are changing the uncertainty by our actions, this brings the model into the realm of decision-dependent uncertainty, discussed in detail in the next section.

3.7 Modelling decision-dependent uncertainty

The scenario-tree-based approach described in Section 3.5 has one problem: its structure is given, so the timing of the uncertainty revelation is fixed. This means that this approach can only handle exogenous uncertainty, independent of the decisions made by the model. While this may work fine for prices or demand, it is not appropriate for technological uncertainty, project costs and duration. In those cases, the revelation of uncertainty depends on when we start the project, not on a fixed time.

To address this issue, we use an approach of 'dynamic' non-anticipativity constraints (Jonsbråten et al., 1998; Goel and Grossmann, 2006). This is illustrated in Fig. 4: assume that we already have a scenario tree for all the 'standard' stochastic parameters, such as the one in Fig. 4a, and now want to introduce a new parameter whose values we learn by starting a specific project. For simplicity, assume that the parameter can take only two values, 'high' and 'low'. To model the situation, we make two copies of the scenario tree, one for each



(a) The original tree

(b) Multiplied tree with non-anticipativity constraints

(c) The same tree with a given solution

Figure 4: Scenario trees with decision-dependent uncertainty, with one uncertain parameter with two possible values. The original scenario tree from (4a) is in (4b) multiplied for the two values of the uncertain parameter (shown in red and blue). The corresponding nodes are connected by non-anticipativity constraints, shown by dashed lines. In (4c), the decision revealing the uncertainty is taken in the highlighted node. Active non-anticipativity constraints are shown with solid lines, inactive with dotted lines.

value, and then connect all corresponding nodes by a constraint requiring that if the project has not been started yet at a node (t, s), then all variables at this node must be equal for the two trees. This ensures that the model cannot distinguish between the 'high' and 'low' states, before we have started the project. In other words, we have non-anticipativity constraints, conditional on not having started the project. This is illustrated graphically in Fig. 4b. Figure 4c then shows a situation for a particular solution where we start the project at the highlighted node of the tree. We see that we are allowed to distinguish between the two cases only in nodes following the decision. In all other nodes, the non-anticipativity constraints are active, so we have to take the same decisions in both cases.

Note that the constraints do not have to be bound to the start of the project; one can just as easily connect them to the end of the project or to any point in between. One can, for example, assume that we are able to distinguish between low, medium and high project costs (or duration) already when we are halfway through the project.

The obvious challenge with this workaround is that it increases the size of the optimization problem dramatically: if the unknown parameter has n_1 states, we have to multiply the whole scenario tree (or at least the part from the earliest possible start time of project $\tilde{p}$ onwards) n_1 times . Including another decision-dependent parameter with n_2 states will again increase the problem n_2 times, so we will then have $n_1 \cdot n_2$ copies of the original tree. It follows that we can use this approach only for a limited number of parameters with a limited number of states.

For this reason, we have implemented this feature in InfraPlan for a group of projects: the constraints are dropped after a project from the group has been started. We have used this approach for the cost of installing a new technology on different parts of the network. Here, the main uncertainty is caused by the fact that it is a new technology, so it is natural to assume that the uncertainty is revealed after one of the projects has been started (and, hence, experience with the new technology has been gained).

4 Case study: Infrastructure upgrade and construction projects on the Dovre and the Gjøvik Line

As a small but realistic case on the Norwegian railway network, we study the Dovre and the Gjøvik Line to the north of Oslo (*osl*), illustrated in Fig. 5. With a variety of projects, partially depending on each other, this illustrates typical analyses the InfraPlan model may be used for.

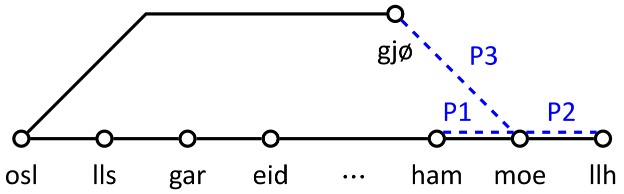


Figure 5: The Dovre and the Gjøvik Line with investment projects P1 (double track Hamar – Moelv), P2 (double track Moelv – Lillehammer) and P3 (connector Gjøvik–Moelv)

Table 1: Projects in the case study

Project	Type	Section	Duration [years]	Latest Start [period]
P1	upgrade	(*ham*) – (*moe*)	3	5
P2	upgrade	(*moe*) – (*llh*)	3	5
P3	new line	(*gjø*) – (*moe*)	4	7

We consider a time horizon of sixty years with ten periods of varying length, five years for the first period, followed by two years for the next five periods to 25 years for the last one:

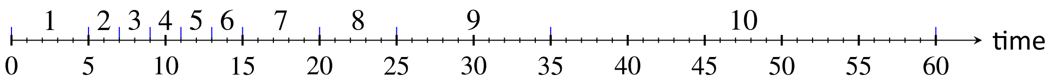


Figure 6: Time periods (above) and optimization horizon in years (below) in the case study

Potential infrastructure upgrade projects on the Dovre Line concern upgrading the whole section between Hamar (*ham*) and Moelv (*moe*) from single to double track (project P1), upgrading to double track between Moelv and Lillehammer (*llh*, P2), and building a connector between the two lines across lake Mjøsa, joining Gjøvik (*gjø*) and Moelv stations (P3). The main project data are presented in Table 1.

Projects P1 and P2 enable some travel time improvements between Oslo and Moelv and between Moelv and Lillehammer, respectively. Once both projects have been carried out, new train schedules can be introduced on the Dovre Line. Completion of P3 enables the introduction of yet another set of services and train schedules on both lines. The new timetable, however, presupposes sufficient capacity on the Dovre Line and, hence, the completion of projects P1 and P2. The new services and schedules thus require all three projects to have finished.

Passenger demand data are based on actual traffic measurements and population growth predictions for the considered time horizon. Obviously, no current demand data are available for the connector and we estimate demand based on concept study reports and other traffic in the area. Other model parameters such as infrastructure or rolling-stock characteristics are, mostly, taken from publicly available sources while socio-economic data are based on the Norwegian Rail Administration's cost-benefit model Merklin and their method handbook for socio-economic analyses (Jernbaneverket, 2011). We could not obtain sufficient data to include freight transport in our analyses.

The model seeks to balance the start of the projects, taking into account their contributions to the expected net benefit of the whole portfolio: those with negative benefit will be started as late as possible (if at all), while those with positive benefits will be initiated early. Discounting intensifies this effect, while precedence or other dependency relations as well as the cost-benefit ratio between dependent projects will also play an important role. Running the InfraPlan model on the situation described above, no project was suggested to be carried out. In other words, the model could not find any combination of project investments over the sixty-year horizon such that the expected net benefit is improved compared to the reference alternative. In particular, even if the improved timetables on the Dovre Line have positive benefits, they do not outweigh the costs of projects P1 and P2, regardless of the timing.

The model set-up makes it easy to analyze related situations and to investigate effects of various measures. For instance, it may be required that a given project has to be carried out during the considered time horizon. Then, the model will determine the best point of time to do so and how this will affect other project options. As an example, assume that all three construction projects P1, P2 and P3 shall be carried out. Each project has a negative cost-benefit balance and will, hence, be started as late as possible, P1 and P2 in period five and P3 in period seven. The new timetables on Dovre Line will be implemented after P1 and P2 have finished, i.e., from period seven. On the other hand, the new timetables taking advantage of the connector (*gjø*) – (*moe*) turned out to have negative benefits and will not be implemented, despite the infrastructure being in place.

The InfraPlan model makes it also easy to redefine the reference alternative and include some of the project options there, removing their costs and benefits from the analysis. This can show the benefit of single projects or groups of projects given these other investments have already been carried out.

As a further alternative, we investigate the effect of periodical budget limits, prohibiting large investments to be carried out in parallel. We keep the requirement that all three projects shalls be implemented. In this case, projects P1 and P3 get initiated one period earlier (P1 in period 4 and P3 in period 6). With slightly more slack in the budget limits during the later, longer periods, P3 may be carried out in period 7 as before, but now the model suggests to start project P2 first, in period 3, and postpone the start of project P1 until period 5, resulting in a somewhat higher net benefit. This effect is mainly due to differences in the duration of both projects and time periods and illustrates that small changes in the model set-up and parameters may affect also apparently independent decisions.

5 Conclusions

In supply chain management and portfolio optimization, coordination of asset acquisitions and other decisions is considered the key to improve margins and to reduce risk. The InfraPlan optimization model brings these concepts into a CBA framework, to be used by analysts for decision support for railway infrastructure investments. It includes elements of network design, line/service planning and demand modelling for a more comprehensive analysis of investment options. In combination with a multi-stage stochastic programming approach to address several forms of uncertainty, this can yield more cost-efficient and robust solutions than current practice allows.

In this report, we provided an overview of the InfraPlan model, addressing modelling aspects dealing with the specifics of railway planning. The strength of the model lies in providing a holistic perspective on the planning process, helping to prioritize a portfolio of projects in relation to each other or to phase in projects over time, taking into account dependencies, budget limits or other economic considerations. It can also point out effects of given project prioritizations or timing decisions under these aspects. By means of a realistic case study from a central Norwegian region, we demonstrated typical situations where the model properties are advantageous over traditional planning approaches.

For models without decision-dependent uncertainty, off-the-shelf solvers find optimal solutions for large-scale real-life model instances within minutes. However, when including decision-dependent uncertainty, the problem size grows rapidly, which makes it hard to address large instances. Future research should, therefore, focus on improving the efficiency of the solution process for such problems.

Another important topic for model improvement concerns projects establishing new infrastructure where none existed before, such as the connector in the case study in Section 4. In this case, no reference alternative exists against which potential solutions can be compared. The calculation of generalized costs would then needs to be more comprehensive, taking into account other modes of transport in the area. However, this would move the model more into the direction of a traffic or transport model rather than a planning-oriented tool. Clearly, more effort is needed to find a well-thought-out approach for this problem.

Once a case has been defined, it is easy to carry out a diversity of analyses—and combinations thereof—and to run the optimization model under slightly changed assumptions. This allows, for example, to test effects of various measures on the performance of the project portfolio or to identify which actions are required to achieve certain solutions. It opens also up for analyses from several perspectives, taking into account and balancing various criteria to find robust decisions.

Investments in railway infrastructure, as well as other transport infrastructure, represent a major part of public spendings. A fair appraisal of investment alternatives is important to ensure both fair comparison between projects and efficient use of public money. In this report, we have presented a model that can consider a portfolio of investment alternatives, thus providing decision-makers with a better basis for making informed decisions.

Acknowledgements

We would like to thank Marius Fossen, Jon-Kristian Ryan Hovland, Gunnar Markussen, Malene Nerland and Per Jorulf Øvervik at Jernbanedirektoratet (Norwegian Railway Directorate) for many useful discussions and constructive feedback during development and testing of the model.

References

K. J. Arrow. The economic implications of learning by doing. *The Review of Economic Studies*, 29(3):155–173, 1962. URL `http://www.jstor.org/stable/2295952`.

R. Bai, S. W. Wallace, J. Li, and A. Y.-L. Chong. Stochastic service network design with rerouting. *Transportation Research Part B: Methodological*, 60:50–65, 2014. doi: 10.1016/j.trb.2013.11.001.

D. Bertsekas. *Dynamic Programming: Deterministic and Stochastic Models*. Prentice Hall, 1987.

J. R. Birge and F. Louveaux. *Introduction to stochastic programming*. Springer, 2nd edition, 2011.

A. Boardman, D. F. Greenberg, A. Vining, and D. Weimer. *Cost-Benefit Analysis: Concepts and Practice*. Prentice Hall, 4th edition, 2014. ISBN 0137002696.

M. Bussieck, T. Winter, and U. Zimmermann. Discrete optimization in public rail transport. *Mathematical Programming*, 79(1–3):415–444, 1997. ISSN 0025-5610. doi: 10.1007/BF02614327.

M. Fodstad, R. Egging, K. Midthun, and A. Tomasgard. Stochastic modeling of natural gas infrastructure development in Europe under demand uncertainty. *The Energy Journal*, 37:5–32, 2016. doi: 10.5547/01956574.37.SI3.mfod.

V. Goel and I. E. Grossmann. A class of stochastic programs with decision dependent uncertainty. *Mathematical Programming*, 108(2–3):355–394, 2006. doi: 10.1007/s10107-006-0715-7.

M. F. Gorman. Evaluating the public investment mix in US freight transportation infrastructure. *Transportation Research Part A: Policy and Practice*, 42(1):1–14, 2008. ISSN 0965-8564. doi: 10.1016/j.tra.2007.06.012.

C. Gundegjerde, I. B. Halvorsen, E. E. Halvorsen-Weare, L. M. Hvattum, and L. M. Nonås. A stochastic fleet size and mix model for maintenance operations at offshore wind farms. *Transportation Research Part C: Emerging Technologies*, 52:74–92, 2015. doi: 10.1016/j.trc.2015.01.005.

L. Hellemo, K. Midthun, A. Tomasgard, and A. Werner. Multi-stage stochastic programming for natural gas infrastructure design with a production perspective. In H. I. Gassmann, S. W. Wallace, and W. T. Ziemba, editors, *Stochastic Programming: Applications in Finance, Energy, Planning and Logistics*, World Scientific Series in Finance. World Scientific, 2013.

Jernbaneverket. Metodehåndbok JD 205. Samfunnsøkonomiske analyser for jernbanen, versjon 3.0 (in Norwegian), 2011. URL `http://hdl.handle.net/11250/279494`.

T. W. Jonsbråten, R. J. B. Wets, and D. L. Woodruff. A class of stochastic programs with decision dependent random elements. *Annals of Operations Research*, 82:83–106, 1998.

G. A. Klutke, P. C. Kiessler, and M. A. Wortman. A critical look at the bathtub curve. *IEEE Transactions on Reliability*, 52(1):125–129, 2003. ISSN 0018-9529. doi: 10.1109/TR.2002.804492.

B.-L. Lin, Z.-M. Wang, L.-J. Ji, Y.-M. Tian, and G.-Q. Zhou. Optimizing the freight train connection service network of a large-scale rail system. *Transportation Research Part B: Methodological*, 46(5):649–667, 2012. doi: 10.1016/j.trb.2011.12.003.

P. Mackie and J. Preston. Twenty-one sources of error and bias in transport project appraisal. *Transport Policy*, 5(1):1–7, 1998. ISSN 0967-070X. doi: 10.1016/S0967-070X(98)00004-3.

T. L. Magnanti and R. T. Wong. Network design and transportation planning: Models and algorithms. *Transportation Science*, 18(1):1–55, 1984. doi: 10.1287/trsc.18.1.1.

T. Odgaard, C. Kelly, and J. Laird. Current practice in project appraisal in Europe. project report HEATCO project. Technical report, University of Leeds, 2006.

N. O. Olsson, A. Økland, and S. B. Halvorsen. Consequences of differences in cost-benefit methodology in railway infrastructure appraisal – A comparison between selected countries. *Transport Policy*, 22:29–35, 2012. ISSN 0967-070X. doi: 10.1016/j.tranpol.2012.03.005.

D. J. Thomas and P. M. Griffin. Coordinated supply chain management. *European Journal of Operational Research*, 94:1–15, 1996.

A. Werner, K. Tolstad Uggen, M. Fodstad, A.-G. Lium, and R. Egging. Stochastic mixed-integer programming for integrated portfolio planning in the LNG supply chain. *The Energy Journal*, 35(1), 2014. doi: 10.5547/01956574.35.1.5.

S. A. Zenios. *Practical financial optimization. Decision making for Engineers*. Blackwell, 2007.

SINTEF

Technology for a better society

www.sintef.no