Modeling uncertainty dynamics in public transport optimization

Alessio Trivella

Francesco Corman

ETH Zürich   **May 2019**

ETH Zürich

Modeling uncertainty dynamics in public transport optimization

Alessio Trivella, Francesco Corman
IVT - Institute for Transport Planning and Systems
ETH Zürich
Stefano-Franscini-Platz 5, 8093 Zürich, CH
phone: +45 52 90 63 39
{alessio.trivella, francesco.corman}@ivt.baug.ethz.ch

May 2019

Abstract

Public transport networks such as bus and railway networks are highly complex systems. In fact, multiple sources of uncertainty including fluctuating passenger demand, variable road and traffic conditions, weather, and technical failures affect the network performance and reliability. These uncontrollable, stochastic factors follow intricate dynamics in space and time that makes it difficult to incorporate them into important decision-making processes of traffic management. For example, understanding how delays evolve (fade out absorbed by available buffer times, remain the same, or propagate through the network) is critical to undertake correct rescheduling actions for vehicles in the presence of delays or disruptions. Moreover, the number of stochastic factors is usually very large due to the many moving units or network links, which poses further modeling challenges. Goal of this paper is twofold. First, we review the existing stochastic models of the uncertainty employed in the public transport optimization literature, underlying their merits and shortcomings. Second, we define a roadmap for modeling high-dimensional uncertainties in public transport networks in a sound manner, with the goal of incorporating this uncertainty into stochastic optimization approaches.

Keywords

Railway and bus networks; uncertainty dynamics; stochastic processes; stochastic optimization.

1 Introduction

Public transport (PT) networks such as bus and railway networks play a crucial role in transportation systems due to their large capacity and environmental benefits. Despite the excellent standards of PT networks in Switzerland, societal and environmental targets focus strongly on increasing modal share and quality of collective transport. This puts increasing pressure on PT companies to improve performance in terms of reliability, frequency, and punctuality. Meanwhile, these companies are being forced to increase their cost efficiency, due to limited investments and funding or even cost reductions.

The variability and delays in operations represent the single largest threat influencing service reliability and making improvements in the efficiency of PT networks complex to achieve. Buses and railways are operated according to plans (timetables) that are determined in advance. When an unexpected disturbance occurs, operations might deviate from this plan and cause delays. Even though Switzerland is one of the countries with the highest realized PT punctuality, delays are not uncommon. For example, more than 10% of train runs experienced more than 3 minutes delay SBB (2019). These disturbances can easily propagate through time and space in bus and railway networks according to complex dynamics that are hard to predict. Besides crew and vehicle plans, delays can significantly affect passengers (e.g., missed connections), strengthening the perception of PT unreliability.

The operations of PT networks is complex also because it relies on the resolution of many interconnected optimization problems, for instance, determining the arrival and departure time of vehicles, rescheduling the system in the presence of delays, and operating vehicles in an energy-efficient manner. One way to improve the performance of PT networks would be by solving these optimization problems with the objective of reducing delays and their propagation in the network. In other words, taking tactical and operational decisions that would make the PT network more efficient, more robust to perturbations, and more reactive to recover delays. This goal could be achieved by developing optimization models that would account for current and possible future uncertainties when making decisions. Most PT literature, however, has so far employed a deterministic optimization approach in which either the system is assumed to be completely known, or expected values are assumed sufficient to determine system parameters. In reality, multiple sources of uncertainty such as fluctuating passenger demand, variable road and traffic conditions, weather, and technical failures, affect the performance and reliability of public transport networks.

Correctly modeling the uncertainty and representing it in a way that is suitable for an optimization model is challenging and requires several modeling steps (Trivella 2018): (i) identifying the

sources of uncertainty affecting decisions, (ii) selecting a model to describe the stochastic evolution of the uncertainty, e.g., a stochastic process, (iii) calibrating the parameter of the model using historical data, and (iv) generating scenarios of the uncertainty form the calibrated model. Thus, it is important that uncertainty in PT networks is accurately modeled to produce realistic scenarios for the future traffic, passenger demand, weather, and delays. The uncertainty related to delays, for example, might involve either infrequent disruptions or small-medium but more frequent delays, which might follow very different dynamics in space and time. After uncertainty is modeled, tactical and operational problems that explicitly account for this uncertainty in the decision making process have to be formulated and solved. Considering "uncertainty-aware" methodology has the potential to significantly improve upon current industry practices and state of research, in which future uncertainties are mostly neglected.

The rest of this paper is organized as follows. We start in Section 2 by giving a brief overview of optimization under uncertainty for the benefit of the reader. To underscore the need of incorporating uncertainty in PT optimization, in Section 3 we perform a literature review on the existing stochastic models of uncertainty in PT networks underlying their shortcomings, with a primary focus on railways and buses. Notice that in this review we are interested in how uncertainty is considered and modeled in order to make decisions, i.e., within optimization models. We do not consider empirical papers that, e.g., calibrate probability density functions (despite there exist several ones). Then, in Section 4 we define a roadmap for modeling high-dimensional uncertainties in PT networks in a sound manner for incorporation into optimization processes. In Section 5, we draw conclusions on uncertainty modeling in PT optimization.

2 Stochastic optimization background

Optimization under uncertainty, or stochastic optimization, refers to a collection of quantitative methods to make better decisions in the presence of uncertainty. While deterministic optimization is handled using a universal mathematical programming framework, stochastic optimization encompasses different modeling techniques and solution approaches including stochastic programming (Birge and Louveaux 2011), robust optimization (Ben-Tal *et al.* 2009), and approximate dynamic programming (Bertsekas 2011, Powell 2011). Tackling a stochastic optimization model is usually significantly more challenging than a deterministic model (which can be itself an $\mathcal{NP}$-hard problem). The reasons behind this additional complexity is often associated with the search over policies (i.e., collections of decision functions) rather than scalars or vectors (Powell 2018) and with the necessity of establishing a good model to describe the stochastic evolution of the uncertainty.

Stochastic optimization approaches are widely recognized and applied in a variety of contexts such as energy operations and planning (Wallace and Fleten 2003, Boomsma *et al.* 2012, Trivella *et al.* 2018), finance (Ziemba and Vickson 2014), and healthcare (Ahmadi-Javid *et al.* 2017), and have led to remarkable achievements. Also, research in air traffic control, which is similar to railway traffic control in terms of capacity and passenger flows (Pellegrini and Rodriguez 2013), has already pioneered stochastic optimization in planning of operations, resulting in major improvements (Glover and Ball 2013, Jacquillat and Odoni 2015). On the other hand, the stochastic optimization literature in railway and bus transportation is far behind and still requires a substantial development. This paper attempts indeed to narrow this knowledge gap by understanding how bus and railway optimization problems can be tackled using the realistic environment dynamics neglected in current deterministic models.

3 Uncertainty models in the public transport optimization literature

The functioning and operations of a PT network require solving a number of interconnected problems at tactical level (e.g., timetabling) and operational level (e.g., traffic control) that in reality are usually affected by some degree of uncertainty. As mentioned in Section 1, taking this uncertainty into account when solving these problems is critical to improve the performance and reliability of PT.

In this section, we present some of the most important problems in railway and bus optimization and show that uncertainty has only been considered limitedly or with too simple, hence unrealistic models in the literature. We review the literature on PT rescheduling in Subsection 3.1, on PT timetabling in Subsection 3.2, and train trajectory optimization in Subsection 3.3.

3.1 Real-time rescheduling problem

In this subsection, we consider the *rescheduling* problem, or real time traffic control, which consists in taking corrective actions in the PT network in order to mitigate the impact of delays or disruptions in the system. Given that both the sources of uncertainty and the rescheduling decisions are very different between railway and bus networks, we split our review in two parts: we consider railways first and then move to buses.

3.1.1 Train rescheduling

Railway operations exploit the available infrastructure capacity according to a predefined timetable, i.e., an operating plan that establishes arrival and departure times of trains. When disturbances in the network occur due to e.g. technical failures or fluctuations in passenger demand and travel time, railway operations might deviate from the timetable. Since railway networks are very constrained systems, delays can easily propagate in space and time in a snowball effect (knock-on), especially in saturated infrastructures (Corman *et al.* 2010a).

Timetables with large buffer times between trains ensure that delays are easily absorbed by the network, but imply lower utilization of the infrastructure that makes railway services less attractive and economical (Kroon *et al.* 2009). This approach is also not viable in practice due to the expected increase of railway passenger demand and consequently capacity utilization, as in case of Switzerland (Zischek 2017). A complementary and more practical way to handle delays is through railway traffic control procedures, i.e., adjusting the timetable in real-time by taking actions such as retiming arrivals and departures, reordering trains, and rerouting trains (Corman and Meng 2014). Choosing the right action is however hard for a human traffic controller due to: (i) the large number of rescheduling options available, and (ii) the variability in the system that complicates determining whether a delay will fade out absorbed by the buffers, remain the same, or propagate. Thus, human controllers can only rely on their past experience in evaluating the effect of real-time schedule updates, or on fixed procedures (Hansen and Pachl 2014).

The academic research tackled the railway traffic control problem by developing techniques of mathematical optimization aimed at finding the most suitable (i.e., mathematically optimal) control action. The models developed in the literature are both macroscopic (large networks with only stations and lines; Tomii *et al.* 2005, Meng and Zhou 2011) and microscopic (small networks with high level of details including individual signals, blocks, and switches; Corman *et al.* 2010b, 2011, 2012, Pellegrini *et al.* 2014, Lamorgese and Mannino 2015).These formulations are usually challenging combinatorial optimization problems. Thus, the common assumption to quickly find a solution to these models is a deterministic and static setting, i.e., with full information about future state of the system. As a result, these control approaches actually ignore the uncertainty and variability that they aim to reduce. The use of uncertainty-aware models in railway optimization is recognized as a promising approach which is slowing being adopted by researchers. However, to our knowledge, only a few papers that we review below deal specifically with uncertainty-aware real-time railway operations.

Meng and Zhou (2011) focus on finding a robust train dispatching under random segment running time and segment capacity breakdown duration. The work is relevant as the authors develop a rolling-scenario approach which is scenario-based, i.e., decisions are made by accounting

for multiple possible outcomes of the uncertainty. In the numerical experiments, however, the authors only consider a one-dimensional uncertainty (a capacity breakdown that follows a Gaussian distribution), which makes the problem significantly easier compared to having (i) multiple uncertainties that (ii) evolve over time. Yin *et al.* (2016) consider the energy-efficient real-time rescheduling problem in a metro system using approximate dynamic programming. The system is affected by uncertainty in passenger demand and is high-dimensional. In fact, passenger demand is modeled for each station of the line as a time-dependent Poisson distribution. However, the random variables are independently distributed over stages and are uncorrelated. In a recent work, Ghasempour and Heydecker (2019) also use approximate dynamic programming for a train rescheduling problem in a single junction. The authors account for uncertainty in dwell times at stations and section running time delays, and model them, respectively, using a Weibull distribution and a Beta distribution. The random variables in this paper are also uncorrelated since they are considered as independent and identically distributed, similar to Yin *et al.* (2016).

3.1.2 Bus rescheduling

Bus networks are also frequently affected by delays. In addition to fluctuations in passenger demand and boarding and alighting time, the travel time variability (e.g., due to traffic conditions and weather) is a major source of uncertainty in bus networks (Kieu *et al.* 2014, Ma *et al.* 2016). This makes the delay and system dynamics significantly different than in railway networks. One of the most common effect of this stochastic system dynamics is the so-called *bus bunching*, or *platooning*. A bus service is indeed considered best when times between successive bus arrivals (headways) are equal. However, variability in the network makes it impossible to maintain equal headways. When a bus is behind schedule, it will stay longer at the next stops because more passengers will be boarding and alighting, and the delay will increase further. At the same time, the amount of passengers in the following bus will decrease and the headway between the two buses becomes smaller and smaller, resulting in buses clustering together, or bunching. When this effect happens, bus operators might use different bus rescheduling strategies to restore the desire headways.

Holding strategies are popular and consist in holding a bus at a station to create a certain headway with the bus in front or behind. Bus holding to equalize headways has been studied by academics for many years. In some recent literature, Bartholdi and Eisenstein (2012) suggested to systematically delay buses at control points, while by Delgado *et al.* (2012) developed and solved a non-linear deterministic mathematical program. Another strategy to address bunching is stop skipping, in which a bus is allowed to skip a stop either completely or partially by only enabling passenger drop (Sun and Hickman 2005). A joint holding and stop skipping

strategy was studied by Sáez *et al.* (2012) using a model predictive controller on a short time horizon. Other options to restore headways include short turning, i.e., the bus turns around before reaching the end point, deadheading, i.e., the bus is allowed to skip some stops when empty (Yu *et al.* 2012), or simply speeding up between stops. Similar to the railway literature, research in bus rescheduling has not accounted for uncertainty explicitly in the decision making process to the best of our knowledge.

A stochastic simulator of a bus network in used in Delgado *et al.* (2012). This simulator assumes that (i) passenger arrivals follow a Poisson distribution with different mean at each stop, and (ii) travel times between two stops follow a log-normal distribution. However, the optimization model developed by the authors to take corrective actions (i.e., bus holding time) is deterministic, i.e., it does not account explicitly for future scenarios of the uncertainty but simply uses the mean of these distributions as reference. The model is then re-optimized every time a bus arrives at a station. The rolling horizon approach by Sáez *et al.* (2012) assumes that passengers arrive at stations following an exponential distribution instead. The rescheduling actions in this paper are also taken based on a single forecast of the system evolution, that is, the variability of the future is not exploited when taking decisions.

3.2 Timetabling problem

The real-time rescheduling possibilities for trains and buses discussed above and their effectiveness in presence of delay depend on the *timetable*. Timetables are determined in the planning phase but influence real-time traffic control and vice-versa, i.e., the two problems are interconnected. Constructing a timetable consists in determining the planned arrival and departure times for units at stops/stations and has been extensively studied in the PT literature (see Parbo *et al.* 2016 for a recent survey, and Robenek *et al.* 2016). Determining PT timetables is computationally hard even in deterministic settings because all departure and arrival times are considered as decision variables, giving rise to complex mathematical programs. Nonetheless, some recent literature has also addressed the problem of incorporating robustness in PT timetable design or planning processes in general (see Fischetti *et al.* 2009, Cacchiani and Toth 2018, Lubsy *et al.* 2018, and references therein), where robustness refers to the ability to absorb or resist to unexpected changes, i.e., to continue operations at some level under disturbances. The drawback of these approaches is that they are static, in the sense that they cannot fully integrate the real-time dynamics and rescheduling decisions into the timetabling design phase. A fully stochastic/robust timetabling model would, for each feasible timetable, solve a second-stage problem to evaluate the real-time operational performance of the current timetabling solution under uncertainty or perturbations. Clearly, this fully stochastic perspective would also increase

the complexity of a possible solution approach.

A simpler way to increase robustness towards real-time operations and prevent delay propagation is by properly allocating the buffer times between possibly conflicting events in a timetable. This can be seen as a subproblem of the full timetabling problem, where decision variables are buffer times to insert between events rather than every departure and arrival time. If we had accurate delay dynamics and uncertainty-aware policies for real-time railway traffic control or bus rescheduling, we could in fact exploit these tools to improve the buffer time allocation in timetabling. Kroon *et al.* (2008) discuss the buffer time allocation problem using a stochastic optimization model that maximizes the robustness against stochastic disturbances. The paper assumes that stochastic disturbances are independent of the current timetable and that the simulation of the model does not include traffic control decisions. The initial (primary) disturbances are modeled as truncated exponential distributions, which are then randomly perturbed during the rest of the horizon. The buffer time allocation is also tackled in Jovanović *et al.* (2017) and a robust allocation is found essentially via a knapsack problem reformulation. Despite the approach in Jovanović *et al.* (2017) is interesting, it also has major limitations since the uncertainty: (i) is only considered a-posteriori to evaluate a solution, and (ii) it follows extremely poor dynamics because an initial delay is drawn from a uniform distribution and is then assumed to propagate in a deterministic manner through the network. Thus, the buffer time allocation literature would benefit from more realistic delay dynamics, and by integrating the uncertainty and real-time control decisions in the allocation process using stochastic optimization techniques.

3.3 Train trajectory optimization

The last problem we consider in this review pertains specifically to railways and is the *train trajectory optimization problem*, that is, determining energy-efficient trajectories for trains driving between two stations while fulfilling the scheduled arrival time and the other various operational constraints. This problem has attracted considerable attention in the recent railway literature (see, e.g., the surveys in Yang *et al.* 2016b and De Martinis and Corman 2018). Improving energy efficiency is indeed one of the most important challenges in modern railway transportation because the energy consumption represents one of the largest operating costs (Railenergy 2016). As discussed in Hansen and Pachl (2014), optimizing the speed profiles of individual trains can lead to potential energy savings in the range of 5–20%, and is therefore an attractive measure for railway companies to reduce energy consumption since it does not require any particular investment or infrastructural updates.

The train trajectory optimization problem has been approached in the operations research

literature using a number of methods that include formulating a mathematical program (e.g., mixed-integer and/or non-linear; Wang *et al.* 2013, Wang and Goverde 2016), formulating a boundary value problem with differential equations (Howlett 2000, Howlett and Pudney 2012), and by using dynamic programming on a graph generated from discretized space, time, and speed points (Ko *et al.* 2004, Haahr *et al.* 2017, Zhou *et al.* 2017). The existing literature has primarily studied the train trajectory optimization problem in a deterministic environment where speed profiles are computed neglecting any uncertainty during the trip. In reality, there are random factors that vary for each journey on the same track or even during the journey. These factors include weather conditions (wind, rain, snow), train load (mass of passengers and goods), traction effort and train resistance, for instance, and can affect the energy consumption and consequently the optimal train speed profile. Accounting for these factors would enable deriving driving controls that could be more energy-efficient because they could adapt to this uncertainty. We were able to find only a few papers that deal with train trajectory optimization in an uncertain, or partially uncertain environment.

Yang *et al.* (2016a) consider an integrated timetabling and speed profile optimization problem under variable train mass at inter-stations, tractive force, braking force, and train resistance. However, leaving train mass aside, the other quantities are actually considered "variable" by the authors simply because they are modeled as a function of speed (which in common in microscopic models), but they are not treated as random variables. The paper shows that accounting for uncertainty in train mass can lead to 3% energy saving compared to a deterministic benchmark. Trivella *et al.* (2019) consider uncertainty in wind as a Weibull distribution and suggest to exploit the knowledge of this information to improve the trajectory based on an updated train resistance. The model however assumes that weather information realizes before the beginning of the journey and does not vary dynamically during the journey, therefore stochastic optimization is not used by the authors.

4 A roadmap for uncertainty modeling in optimization problems

In this section, we start by highlighting and discussing the shortcomings of existing models in the literature in Subsection 4.1. Then, in Subsection 4.2 we indicate our suggestions on how to develop in a sound manner models for the stochastic evolution of uncertainty in PT networks. In Subsection 4.3, we discuss how to incorporate these models into an optimization routine.

4.1 Shortcomings in existing models

Our literature review in Section 3 highlighted that uncertainty is relevant in railway and bus networks since there are numerous papers that consider uncertainty in passenger demand, weather, running times, dwelling times, delay duration, etc. However, our review also points out that the current modeling of uncertainty in PT optimization is lacking, mainly for the following two reasons:

1. There is a lack of optimization models that account for uncertainty explicitly in the decision making process, that is, consider future variability through multiple scenarios of the uncertainty. In fact, most optimization models employ a deterministic optimization approach based on a **perfect information assumption or a single forecast of the uncertainty** (expected value) to make decisions. The expected value of a probability distribution conveys only very limited information on the distribution, which is generally insufficient to make good decision in the presence of uncertainty. Intuitively, this issue might be even more acute in case of multimodal distributions because the expectation can fall between two peaks. These distributions have indeed been examined in the PT empirical literature and, for example, Ma *et al.* 2016 suggested that bus travel times follow a multimodal distribution.

2. The dynamics of uncertainty is usually captured in a poor manner by assuming a single stochastic factor and/or by using **simple independent and uncorrelated random variables** that include the uniform distribution, Gaussian, Weibull, Beta, or exponential, to name a few. No paper in our review employs stochastic processes, which are generally considered more appropriate to model the stochastic evolution of uncertain factors over time, compared to independent random variables. In fact, stochastic processes embed a temporal correlation of random variables. For instance, the use of stochastic processes to model uncertainty is well established and common in the energy optimization and finance literature. Below, we identified other shortcomings related to uncertainty modeling:

   - A single-factor model of uncertainty is generally used rather than multi-factor models that assume the uncertainty to be driven by multiple stochastic factors. This is a big simplification in many optimization models that involve a large number of moving units or network links and where a multi-dimensional model of uncertainty appears more appropriate/realistic.

   - Almost no paper assumes correlation among different random variables. Correlation in practice exists between different sources of uncertainty as well as different random

variables in space and in time (e.g., the evolution of delays, number of passengers).
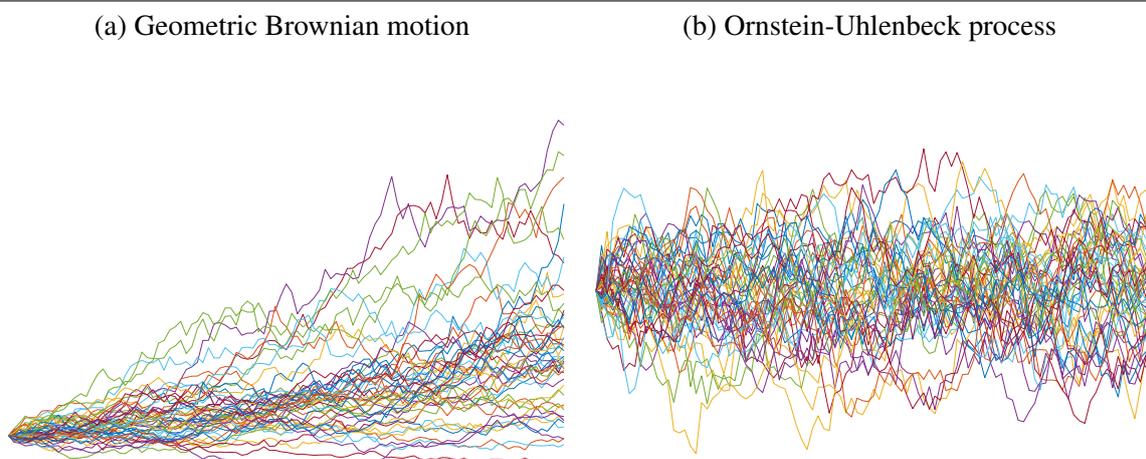
- In some papers only the initial (time-zero) uncertainty is sampled from a probability distribution and is then assumed to evolve deterministically over time. This means that the uncertainty has no dynamics over time and, once the initial value of the uncertainty realizes, then the decision making problem reduces to a deterministic optimization problem.

Given the lack of modeling approaches that consider realistic uncertainty dynamics and the explicit integration of uncertainty into optimization/decision making, in the next sections we propose some guidelines for overcoming these limitations.

4.2 Modeling phase

First, we propose the **use of stochastic processes** to describe the stochastic evolution of the unknown system over time (e.g., passenger demand or road traffic conditions), rather than independent and uncorrelated random variables. Example of stochastic processes are the well-known geometric Brownian motion, Ornstein-Uhlenbeck, and Poisson diffusion, to name a few (Pinsky and Karlin 2010). Each process is able to capture different aspects of the uncertainty dynamics, as it can also be noticed in Figure 1 where we draw sample paths from a geometric Brownian motion (Figure 1(a)) and an Ornstein-Uhlenbeck process (Figure 1(b)). For instance, the former process can capture a trend/drift in the data while the latter process can capture mean reversion. Thus, the stochastic process must be chosen carefully based on the characteristics of the historical data.

Figure 1: Example of stochastic processes

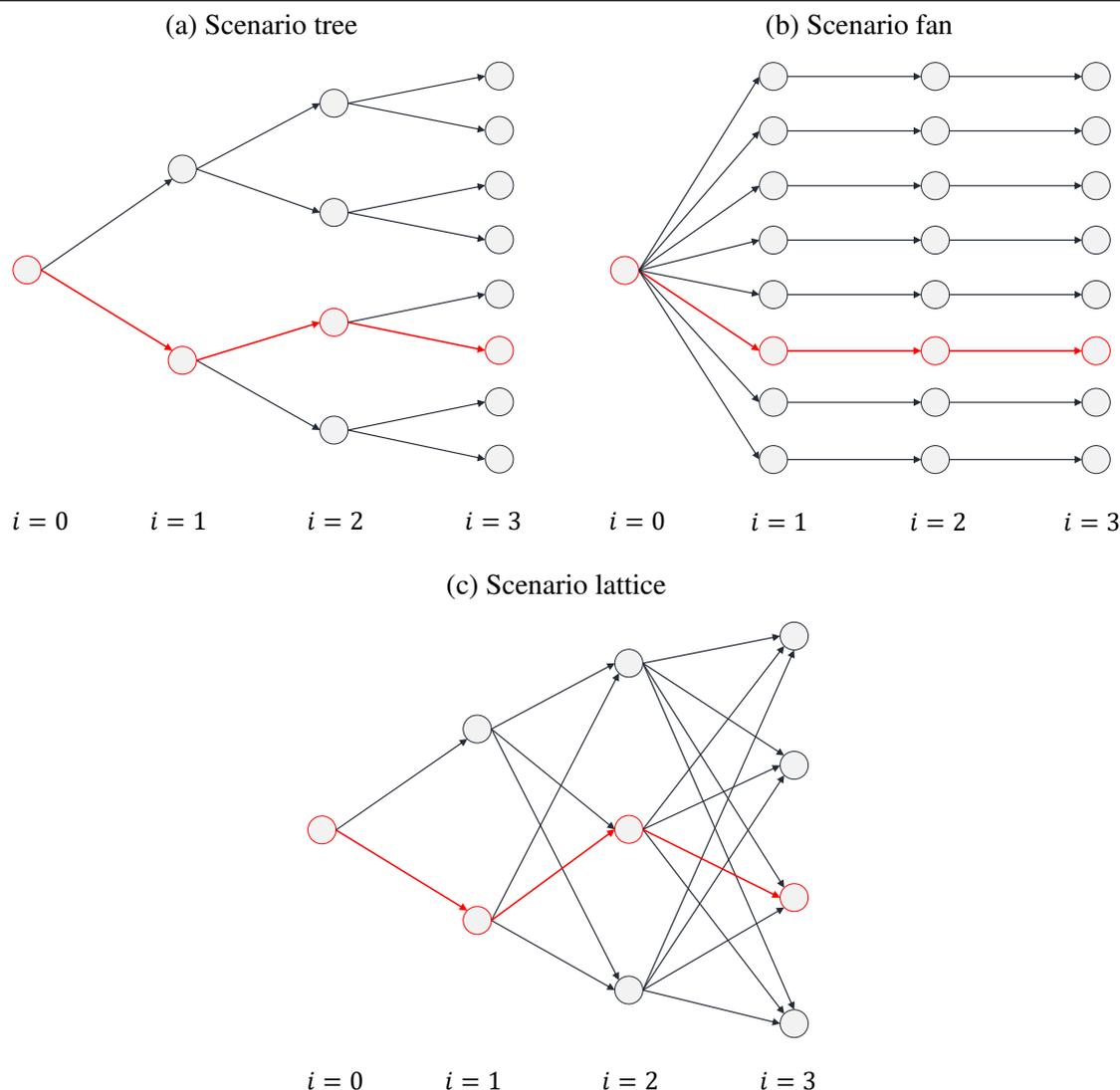| (a) Geometric Brownian motion | (b) Ornstein-Uhlenbeck process |

There exist also more complex processes in which the evolution of an uncertain quantity is driven by multiple stochastic factors, i.e., **multi-factor processes**. A realistic modeling of the uncertainty in a PT network in space and time entails associating each moving unit or link in the network with at least one stochastic factor (i.e., amount of traffic on the link) representing how uncertainty evolves in time in some spatial coordinates (which can either be "Eulerian", i.e. fixed, or "Lagrangian", i.e., following the movement of the vehicle). A large number of moving units or network links results in a high-dimensional dynamical model of the uncertainty, i.e., composed of many stochastic factors. These factors will also have some degree of correlation. For instance, adjacent roads have similar traffic conditions with high probability.

Once a suitable stochastic process has been identified, the next step involves **calibrating the process on real data** to obtain the parameters (e.g. volatility, drift, correlation) that best fit the historical data about passenger demand, traffic, and potentially other uncertainties. For example, the calibration of a stochastic process can be done by applying a Kalman filtering technique (see Hamilton 1995, Schwartz and Smith 2000).

At this point, the calibrated stochastic process can be used to **generate scenarios for the uncertainty** using Monte Carlo simulation. Depending on the structure of the underlying decision making problem, the evolution of the stochastic process can be approximated by a discrete-state scenario process using different representations.

In Figure 2, we show three different ways of representing the uncertainty with discrete scenarios over a four-period horizon: a *scenario tree* (Figure 2(a)), a *scenario fan* (Figure 2(b)), and a *scenario lattice* (Figure 2(c)). In all three cases, a node w_i in the representation denotes a stochastic process outcome at a stage i. In the scenario tree and lattice, given w_i, the distribution of w_{i+1} in the next-period is characterized by finitely many possible outcomes corresponding to the branches exiting w_i and probabilities associated with each branch. A scenario is a path from the root (stage $i = 0$) to a leaf (stage $i = 3$ in the figure) as the ones marked in red. The number of nodes and scenarios in the scenario tree grows exponentially with the number of time periods, which makes handling scenario trees hard in problems with many decision stages. In the scenario lattice, the number of nodes only grows linearly but it might be anyway difficult to handle this structure since there is no unique predecessor for a given node w_i. In contrast, a scenario fan consists of a set of sample paths of the uncertainty generated in Monte Carlo simulation. Scenario reduction techniques are frequently used to decrease the number of scenarios and increase computational tractability while preserving most of the stochastic information enclosed in the original set of scenarios.
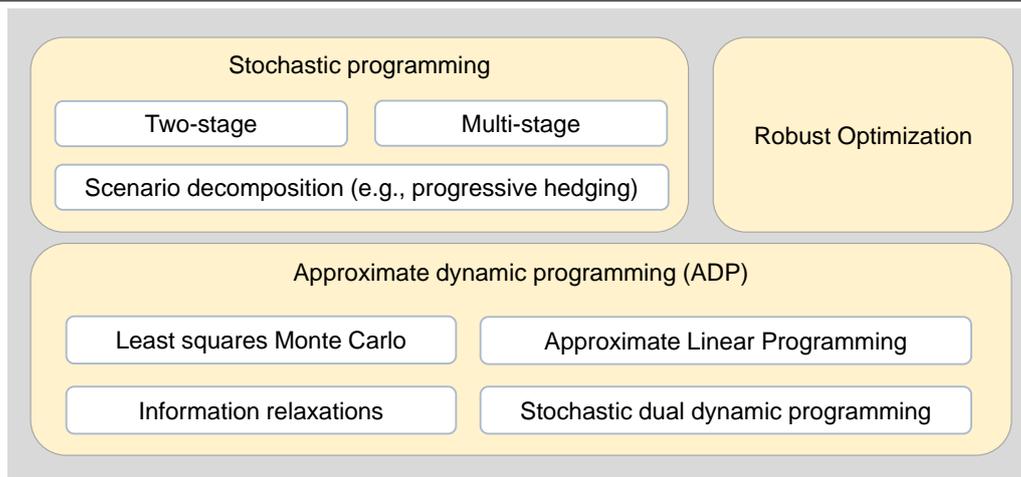
Figure 2: Uncertainty representation

(a) Scenario tree

(b) Scenario fan

$i = 0$ $i = 1$ $i = 2$ $i = 3$ $i = 0$ $i = 1$ $i = 2$ $i = 3$

(c) Scenario lattice

$i = 0$ $i = 1$ $i = 2$ $i = 3$

Source: Trivella (2018).

4.3 Optimization phase

Assume now that a calibrated model for generating scenarios of the uncertainty is in place. The next step in our roadmap consists in formulating a stochastic optimization model to **account for future scenarios of the uncertainty explicitly in the decision making process**. As mentioned, stochastic optimization has the potential to improve (in some case dramatically) the performance of decisions over classic/deterministic optimization. However, it also entails a number of challenges that make formulating stochastic optimization problems more complex. In Figure 3, we display some of the most popular classes of stochastic optimization models. The formulation choice depends on many factors such as the number of stages (e.g., two vs. many),

the dimensionality of the problems (e.g., in the state space, action space, and/or uncertainty outcome space), and the objective function (expected value vs. worst case vs. risk measures). Designing the objective function might be non trivial. Consider for example a railway timetable solution in which a single train has a large delay with large probability and a solution in which many trains have small delays with large probability, so that the total expected delay is the same in both solutions. Which solution is preferable for the operator and for the passengers?

Figure 3: Some popular classes of stochastic optimization models and algorithm.

Finally, we must be able to solve our stochastic optimization model to obtain a solution (i.e., a decision or a control policy) that perform well in the presence of the uncertainty, which means, depending on the objective function, performing well on average (i.e., in expectation) or that optimizes the worst case or some risk measure of the uncertain reward or cost. The classes of models in Figure 3 admit several algorithms to approximate the problem and obtain computational tractability. For example, approximate dynamic programming is an umbrella of methods that include the popular least squares Monte Carlo method (Longstaff and Schwartz 2001, Nadarajah *et al.* 2017, Trivella *et al.* 2018), approximate linear programming (De Farias and Van Roy 2003, Nadarajah *et al.* 2015), stochastic dual dynamic programming (Pereira and Pinto 1991, Löhndorf *et al.* 2013), and information relaxations (Brown *et al.* 2010, Mohseni-Taheri *et al.* 2018) among others, and that can overcome the different curses of dimensionality arising when embedding stochasticity in the problem.

5 Conclusion

In this paper we have discussed the modeling of uncertainty in public transport optimization with a focus on railways and buses. We have reviewed the existing literature by covering

some of the most important optimization problems in public transport including timetabling, real time rescheduling, and energy efficient driving operations. From the extant literature, we identified a lack of modeling approaches that consider realistic uncertainty dynamics and the explicit integration of uncertainty into optimization procedures. In fact, the literature has so far mostly employed poor models for the stochastic dynamics of the uncertainty (e.g., simple, low-dimensional, uncorrelated probability distributions) and a deterministic optimization perspective that either uses perfect information of the future or assumes a single expected value of the uncertainty is sufficient to make decisions.

We therefore provided suggestions on how to overcome these shortcomings from the PT literature in order to improve decision making. Our roadmap for realistic uncertainty modeling in PT optimization consists in two main steps: (i) modeling the evolution of the uncertainty using stochastic processes calibrated based on real data, and (ii) using stochastic optimization techniques that can exploit multiple scenarios of the uncertainty to take better operational and tactical decisions (e.g., stochastic programming, robust optimization, and approximate dynamic programming). Although similar steps are commonly used, e.g., in the energy and finance literature, the PT literature has so far benefited only limitedly from the use of stochastic processes and stochastic optimization methods. Thus, we recommend the adoption of these techniques in future research on PT optimization.

6 References

Ahmadi-Javid, A., Z. Jalali and K. J. Klassen (2017) Outpatient appointment systems in healthcare: A review of optimization studies, *European Journal of Operational Research*, **258** (1) 3–34.

Bartholdi, J. J. and D. D. Eisenstein (2012) A self-coordinating bus route to resist bus bunching, *Transportation Research Part B: Methodological*, **46** (4) 481–491.

Ben-Tal, A., L. El Ghaoui and A. Nemirovski (2009) *Robust optimization*, Princeton University Press.

Bertsekas, D. (2011) *Dynamic Programming and Optimal Control*, Athena Scientific, Belmont, MA, USA.

Birge, J. and F. Louveaux (2011) *Introduction to stochastic programming*, Springer Science & Business Media.

Boomsma, T. K., N. Meade and S.-E. Fleten (2012) Renewable energy investments under different support schemes: A real options approach, *European Journal of Operational Research*, **220** (1) 225–237.

Brown, D. B., J. E. Smith and P. Sun (2010) Information relaxations and duality in stochastic dynamic programs, *Operations Research*, **58** (4) 785–801.

Cacchiani, V. and P. Toth (2018) Robust train timetabling, *In Handbook of Optimization in the Railway Industry*, **258**, 93–115.

Corman, F., A. D'Ariano and I. Hansen (2010a) Disruption handling in large railway networks, *WIT Transactions on the Built Environment*, **114** (1) 629–640.

Corman, F., A. D'Ariano, I. Hansen and D. Pacciarelli (2011) Optimal multi-class rescheduling of railway traffic, *Journal of Rail Transport Planning Management*, **1** (1) 14–24.

Corman, F., A. D'Ariano, D. Pacciarelli and M. Pranzo (2010b) A tabu search algorithm for rerouting trains during rail operations, *Transportation Research Part B: Methodological*, **44** (1) 175–192.

Corman, F., A. D'Ariano, D. Pacciarelli and M. Pranzo (2012) Evaluation of green wave policy in real-time railway traffic managements, *Transportation Research Part C: Emerging Technologies*, **17** (6) 607–616.

Corman, F. and L. Meng (2014) A review of online dynamic models and algorithms for railway traffic control, *IEEE Transactions on Intelligent Transportation System*, **16** (3) 1274–1281.

De Farias, D. P. and B. Van Roy (2003) The linear programming approach to approximate dynamic programming, *Operations Research*, **51** (6) 850–865.

De Martinis, V. and F. Corman (2018) Data-driven perspectives for energy efficient operations in railway systems: Current practices and future opportunities, *Transportation Research Part C: Emerging Technologies*, **95**, 679–697.

Delgado, F., J. C. Munoz and R. Giesen (2012) How much can holding and/or limiting boarding improve transit performance?, *Transportation Research Part B: Methodological*, **46** (9) 1202–1217.

Fischetti, M., D. Salvagnin and A. Zanette (2009) Fast approaches to improve the robustness of a railway timetable, *Transportation Science*, **43** (3) 321–335.

Ghasempour, T. and B. Heydecker (2019) Adaptive railway traffic control using approximate dynamic programming, *Transportation Research Part C: Emerging Technologies*. Articles in advance.

Glover, C. and M. Ball (2013) Stochastic optimization models for ground delay program planning with equity-efficiency tradeoffs, *Transportation Research Part C: Emerging Technologies*, **33**, 196–202.

Haahr, J. T., D. Pisinger and M. Sabbaghian (2017) A dynamic programming approach for optimizing train speed profiles with speed restrictions and passage points, *Transportation Research Part C: Emerging Technologies*, **99**, 167–182.

Hamilton, J. D. (1995) Time Series Analysis, *Economic Theory. II, Princeton University Press, USA*, 625–630.

Hansen, I. A. and J. Pachl (2014) *Railway Timetabling & Operations*, Eurailpress, Hamburg.

Howlett, P. (2000) The optimal control of a train, *Ann. Oper. Res.*, **98** (1-4) 65–87.

Howlett, P. G. and P. J. Pudney (2012) *Energy-Efficient Train Control*, Springer Science & Business Media.

Jacquillat, A. and A. Odoni (2015) An integrated scheduling and operations approach to airport congestion mitigation, *Operations Research*, **63** (6) 1390–1410.

Jovanović, P., P. Kecman, N. Bojović and D. Mandić (2017) Optimal allocation of buffer times to increase train schedule robustness, *European Journal of Operational Research*, **256** (1) 44–54.

Kieu, L.-M., A. Bhaskar and E. Chung (2014) Public transport travel-time variability definitions and monitoring, *Journal of Transportation Engineering*, **141** (1) 04014068.

Ko, H., T. Koseki and M. Miyatake (2004) Application of dynamic programming to the optimization of the running profile of a train, paper presented at the *Computers in Railway Six*, 103–112.

Kroon, L., D. Huisman, E. Abbink, P. Fioole, M. Fischetti, G. Maroti, A. Schrijver and A. Steenbeek (2009) The new Dutch timetable: The OR revolution, *Interfaces*, **39** (1) 6–17.

Kroon, L., G. Maróti, M. R. Helmrich, M. Vromans and R. Dekker (2008) Stochastic improvement of cyclic railway timetables, *Transportation Research Part B: Methodological*, **42** (6) 553–570.

Lamorgese, L. and C. Mannino (2015) An exact decomposition approach for the real-time train dispatching problem, *Operations Research*, **63** (1) 48–64.

Löhndorf, N., D. Wozabal and S. Minner (2013) Optimizing trading decisions for hydro storage systems using approximate dual dynamic programming, *Operations Research*, **61** (4) 810–823.

Longstaff, F. A. and E. S. Schwartz (2001) Valuing American options by simulation: A least-squares approach, *Review of Financial Studies*, **14** (1) 113–147.

Lubsy, M., J. Larsen and S. Bull (2018) A survey on robustness in railway planning, *European Journal of Operational Research*, **266** (1) 1–15.

Ma, Z., L. Ferreira, M. Mesbah and S. Zhu (2016) Modeling distributions of travel time variability for bus operations, *Journal of Advanced Transportation*, **50** (1) 6–24.

Meng, L. and X. Zhou (2011) Robust single-track train dispatching model under a dynamic and stochastic environment: A scenario-based rolling horizon approach, *Transportation Research Part B: Methodological*, **45** (7) 1080–1102.

Mohseni-Taheri, D., A. Trivella and S. Nadarajah (2018) Meeting corporate renewable power targets, `https://ssrn.com/abstract=3294724`.

Nadarajah, S., F. Margot and N. Secomandi (2015) Relaxations of approximate linear programs for the real option management of commodity storage, *Management Science*, **61** (12) 3054–3076.

Nadarajah, S., F. Margot and N. Secomandi (2017) Comparison of least squares Monte Carlo methods with applications to energy real options, *European Journal of Operational Research*, **256** (1) 196–204.

Parbo, J., O. A. Nielsen and C. G. Prato (2016) Passenger perspectives in railway timetabling: A literature review, *Transport Reviews*, **36** (4) 500–526.

Pellegrini, P., G. Marliere and J. Rodriguez (2014) Optimal train routing and scheduling for managing traffic perturbations in complex junctions, *Transportation Research Part B: Methodological*, **59** (1) 58–80.

Pellegrini, P. and J. Rodriguez (2013) Single European sky and single European railway area: A system level analysis of air and rail transportation, *Transportation Research Part A: Policy and Practice*, **57**, 64–86.

Pereira, M. V. F. and L. M. V. G. Pinto (1991) Multi-stage stochastic optimization applied to energy planning, *Mathematical Programming*, **52** (1-3) 359–375.

Pinsky, M. and S. Karlin (2010) *An introduction to stochastic modeling*, Academic press.

Powell, W. (2011) *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, Second ed. John Wiley & Sons, Hoboken, NJ, USA.

Powell, W. (2018) *Stochastic Optimization and Learning: A Unified Framework*, Draft version of January 7, 2018. John Wiley & Sons, Hoboken, NJ, USA.

Railenergy (2016) Innovative integrated energy efficiency solutions for railway rolling stock, rail infrastructure and train operation, `http://www.railenergy.org/background.shtml`.

Robenek, T., Y. Maknoon, S. Azadeh, J. Chen and M. Bierlaire (2016) Passenger centric train timetabling problem, *Transportation Research Part B: Methodological*, **89** (1) 107–126.

Sáez, D., C. Cortés, F. Milla, A. Nunez, A. Tirachini and R. M. (2012) Hybrid predictive control strategy for a public transport system with uncertain demand, *Transportmetrica*, **8** (1) 61–86.

SBB (2019) Facts and figures, `https://company.sbb.ch/en/the-company/profile/facts-and-figures.html`.

Schwartz, E. and J. E. Smith (2000) Short-term variations and long-term dynamics in commodity prices, *Management Science*, **46** (7) 893–911.

Sun, A. and M. Hickman (2005) The real-time stop-skipping problem, *Journal of Intelligent Transportation Systems*, **9** (2) 91–109.

Tomii, N., T. Yoshiaki, N. Tanabe, H. Chikara and M. Kunimitsu (2005) Train rescheduling algorithm which minimizes passengers' dissatisfaction, *Lecture Notes in Computer Science*, **3533** (1).

Trivella, A. (2018) *Decision Making under Uncertainty in Sustainable Energy Operations and Investments*, Ph.D. Thesis, Technical University of Denmark, Kgs. Lyngby.

Trivella, A., S. Nadarajah, S. E. Fleten, D. Mazieres and D. Pisinger (2018) Managing shutdown decisions in merchant commodity and energy production: A social commerce perspective, `https://ssrn.com/abstract=3034869`.

Trivella, A., P. Wang and F. Corman (2019) Train trajectory optimization with external factors: The example of wind, *submitted.*

Wallace, S. and S.-E. Fleten (2003) Stochastic programming models in energy, *Handbooks in operations research and management science*, **10**, 637–677.

Wang, P. and R. M. P. Goverde (2016) Multiple-phase train trajectory optimization with signalling and operational constraints, *Transportation Research Part C: Emerging Technologies*, **69**, 255–275.

Wang, Y., B. De Schutter, T. J. van den Boom and B. Ning (2013) Optimal trajectory planning for trains–a pseudospectral method and a mixed integer linear programming approach, *Transportation Research Part C: Emerging Technologies*, **29**, 97–114.

Yang, X., A. Chen, B. Ning and T. Tang (2016a) A stochastic model for the integrated optimization on metro timetable and speed profile with uncertain train mass, *Transportation Research Part B: Methodological*, **91** (1) 424–445.

Yang, X., X. Li, B. Ning and T. Tang (2016b) A survey on energy-efficient train operation for urban rail transit, *IEEE Trans. Intell. Transp. Syst.*, **17** (1) 2–13.

Yin, J., T. Tang, L. Yang, Z. Gao and B. Ran (2016) Energy-efficient metro train rescheduling with uncertain time-variant passenger demands: An approximate dynamic programming approach, *Transportation Research Part B: Methodological*, **91** (1) 178–210.

Yu, B., Z. Yang and S. Li (2012) Real-time partway deadheading strategy based on transit service reliability assessment, *Transportation Research Part A: Policy and Practice*, **46** (8) 1265–1279.

Zhou, L., L. C. Tong, J. Chen, J. Tang and X. Zhou (2017) Joint optimization of high-speed train timetables and speed profiles: A unified modeling approach using space-time-speed grid networks, *Transportation Research Part B: Methodological*, **97**, 157–181.

Ziemba, W. T. and R. G. Vickson (2014) *Stochastic optimization models in finance*, Academic Press.

Zischek, Y. (2017) The swiss approach to traffic management optimisation, *Global Railway Review*, **3** (1) 637–677.