ABSTRACT

Title of Dissertation:	SELECTION AND SCHEDULING OF INTERRELATED NETWORK IMPROVEMENT PROJECTS UNDER UNCERTAINTIES
	Fei Wu, Doctor of Philosophy, 2023
Dissertation Directed By:	Professor Paul M. Schonfeld, Department of Civil and Environmental Engineering

The analysis of improvements in transportation networks is complicated by the interdependence of those improvements. Changes in links or nodes tend to shift traffic flows. Hence, the benefits of changes in any network component depend on what changes are made at what time in other components. There may also be synergies in the costs of implementing network changes. Methods are needed for selecting and scheduling interrelated network changes under uncertainties regarding demand, costs, implementation times and other factors. The proposed research focuses on optimizing the selection and schedule of interrelated network projects for enhancing a network's performance under various uncertainties. Two multi-level models are formulated for analyzing problems on two types of transportation networks: rail freight networks in Problem 1 and road networks in Problem 2. For rail freight networks, the proposed tri-level model jointly optimizes short-term post-disruption restoration schedules and long-term network development schedules. Its lower level assigns capacitated freight flows to minimize total hourly cost, and its middle level optimizes the restoration sequence for the minimized cumulative cost increment (excess) during

the restoration process under a given disruption scenario. At the upper level, given probabilistic disruption scenarios, network improvement projects are selected from a given set and sequenced to minimize the sum of construction cost and cumulative expected excess over the planning horizon. For road networks, the lower level of the proposed bi-level model performs user-equilibrium (UE) traffic assignment using the Frank-Wolfe (F-W) algorithm. The upper-level model first generates multiple scenarios with samples from the multivariate distribution of multiple correlated uncertain parameters. For a given long-term network improvement plan, an expected present value (PV) of cumulative system travel time cost over the planning horizon plus construction costs of implemented projects is obtained after computing this discounted sum separately under each generated scenario. To minimize this expected PV, the upper level optimizes the improvement plan with a specified selection and sequence of projects.

In both problems, any sequence of restoration or improvement actions is mapped to a unique schedule under the rules based on binding constraints of resources, budget, and required work time. The planning horizon is segmented into short sub-periods based on the improvement schedule to approximate the effects of demand growth and cost discounting. A genetic algorithm (GA) with its operators is customized for optimizing restorations and improvement plans. For road networks, with the internal budget supply based on a fraction of the travel time cost, a set of methods is proposed for determining budget-ready times of projects. The model also allows the use of buses as a mode competing with cars in road networks, and the iteration of mode shares is integrated with the lower-level F-W traffic assignment.

In Problem 1, the proposed model is demonstrated with short-term and long-term numerical cases in a small demand-loaded network. In the short-term problem, the optimized restoration sequences with different numbers of available work teams are obtained by the GA, whose solutions

to relatively small problems are shown to be globally optimal through exhaustive enumeration. The restoration itineraries and schedules of work teams along with corresponding changes in hourly cost are shown in figures. For a small long-term problem, the selection, sequence, and schedule of projects are optimized by exhaustive enumeration. Sensitivity analyses show that the availability of more work teams greatly reduces cumulative expected excess and thus justifies implementing fewer improvement projects. With a longer planning horizon, lower construction costs, and a higher interest rate, more improvements are favored. Evaluations of cumulative expected excess with probabilistic growth rates of demand are provided. The model is also tested on a larger network where the long-term improvement plan is optimized by the GA. The quality of its solution is statistically verified. In Problem 2, 50 scenarios are generated by a quasi-Monte-Carlo sampling method. Demand growth rate, external budget supply, and the multiplier of required construction time are three correlated uncertainties whose effects are analyzed. The selection and sequencing of improvement projects are optimized by the GA. When introducing buses in the network, two operation modes are considered: using dedicated bus lanes and sharing roads with cars. The iteration of bus shares shows that the bus operation is more desirable with dedicated lanes, and that the potential demand from bus passengers greatly affects that desirability. In sensitivity analyses the factors favoring the implementation of projects include a higher demand level, a higher value of travel time, lower construction costs, and a higher demand growth rate. Across the scenarios, the demand growth rate and the external budget supply directly affect the PVC and last completion time (LCT), respectively. Modifying correlation coefficients of uncertain parameters has slight impact on the minimized PVC, but greatly affects the correlation between PVC and LCT for a given improvement sequence.

SELECTION AND SCHEDULING OF INTERRELATED NETWORK IMPROVEMENT PROJECTS UNDER UNCERTAINTIES

by

Fei Wu

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Professor Paul M. Schonfeld, Chair
Professor Cinzia Cirillo
Professor Xianfeng Yang
Professor Bilal M. Ayyub
Professor Michael C. Fu

Acknowledgements

First and foremost, I express my sincere gratitude to Dr. Paul Schonfeld, the advisor of my graduate studies and the chair of my dissertation committee, for all his persistent, affectionate technical assistance on my dissertation as well as other projects I have worked on over the past five years. I am grateful to other committee members – Drs. Xianfeng Yang, Bilal Ayyub, Cinzia Cirillo, Ali Haghani, and Michael Fu – for taking their time to participate in my proposal and dissertation defenses, review my work, and provide insightful suggestions. I am also thankful to other professors who have taught instrumental courses in transportation engineering and its relevant state-of-the-art fields.

I would like to express my appreciation for the support from the University Mobility and Equity Center, led by Morgan State University, and the Federal Railroad Administration, which partially funded the research in this dissertation. I acknowledge the University of Maryland supercomputing resources (http://hpcc.umd.edu) made available for conducting the research in this dissertation.

I also wish to thank my family members in China, who have been constantly understanding and supporting my doctoral study. I owe my gratitude to my roommates and CEE peers for discussing technical topics, having fun together amid work, and providing me confidence to overcome challenges. Finally, I would like to pay my special thanks to Minga, the adorable cat. Her company over the past four years has been keeping me feel safe and sound despite all the uncertainties in the world.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BPR	Bureau of Public Roads
CPU	Central Processing Unit
CV	Coefficient of Variation
F-W	Frank-Wolfe
GA	Genetic Algorithm
LCT	Last Completion Time (of projects)
LP	Linear Programming
OD	Origin-Destination
OF	Objective Function
PMX	Partial Mapped Crossover
PV	Present Value
PVC	Present Value of Cost
TC	Total Cost
UE	User Equilibrium

Chapter 1: Introduction

1.1 Background

In various types of transportation network, freight, passengers, and vehicles move through arcs (links) and nodes under various constraints to satisfy demands from users over the network. For example, in a rail freight network, commodities are shipped under constraints such as on railroad capacities, arrival and departure time windows, and limited numbers of cars and locomotives, to satisfy shipping demands between all origin-destination (OD) pairs. Capacitated freight flows between OD pairs can be optimally assigned to different paths to minimize the overall cost that includes supplier's shipment cost and users' waiting time cost for the arrival of requested items. In a highway network, with travel demands from different origins to different destinations over the network, travelers (vehicles) move along links (roads or lanes), and the travel time and speed through each link is directly affected by its unit-time traffic flow with respect to its capacity. Traffic flows can be assigned in a pattern where no individual traveler (vehicle) between an OD pair can reduce its OD travel time by switching its path, which is defined as user equilibrium (UE).

Occasionally, network operations are disrupted by various events. After occurrence of a disruption event, capacities of multiple nodes and links in the network are reduced (sometimes to zero), while travel times through some links may be greatly increased. In a rail freight network, these changes lead to increases in overall cost as well as in freight that cannot be shipped immediately by rail, even with optimal assignment of capacitated freight flows. For example, Woodburn (2019) empirically analyzed an

unplanned 53-day closure of part of the British rail freight network in 2016. The results of the analysis showed fewer freight trains, extensions of scheduled journey times, and worsened train punctuality as evident consequences of this disruption, despite slow improvement of service over time. The expected closure duration was first extended but later shortened, due to uncertainties over renewal of a damaged bridge.

Multiple alternative restoration plans for the recovery of network operations are usually feasible, given that a limited number of components can undergo restoration concurrently under various resource constraints. Each restoration step takes some time, and completion of each step yields a new network state with updated capacities of components. For reducing excessive costs, restoration may be more urgent for some components than others. The problem for network operators is to find the best sequence and schedule for restoring damaged components so that the cumulative increment of abnormal cost (defined as "excess" here) during the restoration process is minimized. This optimization in response to a disruption scenario is a short-term decision, while long-term improvement of a transportation network considers a range of possible disruption scenarios. In a certain time interval, with given occurrence probabilities of disruption scenarios, the expected value can be obtained based on the minimized excess values in the short-term restoration process. This expected value can be treated as a measure of network resilience against uncertain disruptions. When deciding on a long-term network improvement plan over a planning horizon, multiple projects are available, each having its own requirements for labor, equipment, materials, budgets, and construction durations. Implementation and completion of an improvement project reduces the expected excess mainly by increasing capacities of network components

on adding new links into the network. One problem for network managers is to select and schedule projects to minimize the sum of cumulative expected excess over the planning horizon and costs of the projects to be implemented.

In fact, both short-term restoration and long-term improvement of a transportation network involve selection, sequencing and scheduling of interrelated projects. Multiple projects can be regarded as "interrelated" if the effects of individually completing these projects on network performance cannot be simply aggregated for obtaining the effect of completing them together. To capture this interrelation, measures of network performances (e.g., total travel time, total shipment cost, and total fare collected) need to be evaluated individually for a network configuration after completing specific projects. Moreover, the long-term decisions on network improvement can be optimized for the objective of better network resilience or robustness against uncertainties – regarding not only the disruptions, but also the changes in demand levels, the degradation of facilities, and the expense of the candidate improvement projects, among other factors. In the real world, multiple uncertain factors interact, which should motivate planners to consider correlated uncertainties when optimizing project selection and scheduling.

1.2 Scope of study

This doctoral dissertation focuses on optimizing the selection, sequencing, and scheduling of interrelated network restoration and improvement projects under various uncertainties, such as disruptions, demand, budget supply, and implementation times. Different problems are discussed for two types of transportation networks: rail freight networks and road networks.

1.2.1 Problem statement and objectives for rail freight networks

For rail freight networks, an integrated tri-level model is proposed for jointly optimizing short-term post-disruption restoration schedules and long-term network development selection and schedules. At the lower level, under given settings of demand and component capacity, capacitated freight flows are assigned using linear programming (LP) to minimize total hourly cost. At the middle level, given a disruption scenario, the objective is to find the restoration sequence that minimizes the cumulative cost increment (excess) incurred during the restoration process. At the upper level, given probabilistic disruption scenarios, network improvement projects are selected from a given set and sequenced to minimize the sum of construction cost and cumulative expected excess over the planning horizon. Any sequence of restoration and improvement actions is mapped to a unique schedule consisting of starting and ending times of projects under rules based on binding constraints regarding resource, budget, and required construction time.

1.2.2 Problem statement and objectives for road networks

For road networks, a bi-level model is proposed for optimizing the selection and schedule of network improvement projects. The lower level performs UE traffic assignment for a certain configuration of link capacities and traffic demands. The upper level, at first, generates multiple scenarios with samples from three correlated uncertain parameters – demand growth rate, external budget supply, and the multiplier of required construction time of projects. In the computation under each generated scenario, each improvement sequence being evaluated is mapped to a unique schedule based on binding constraints of budget and construction time. Under this schedule the

planning horizon is segmented into multiple phases with time-varying network configurations, and the system cost of travel time is accumulated over the planning horizon. For a given improvement sequence, after computing the present value (PV) of the sum of cumulative travel time cost plus construction costs of implemented projects under all the generated scenarios, an expected sum can be obtained. To minimize this expected sum, the upper level seeks for the optimal selection and sequence of road improvement projects.

Chapter 2: Literature review

2.1 Transportation system resilience

In a review on vulnerability and resilience of transportation system, Mattsson and Jenelius (2015) summarized that evaluation of disruption-related systematic impacts on the demand and supply sides, along with the application of graph theory, forms the basis of mainstream analysis approaches in this field. Gu et al. (2020) reviewed similarities and differences among reliability, vulnerability, and resilience, which are the three commonly used concepts in evaluating network performance under disruption. They revealed that these concepts are distinctive in terms of emphasis, sensitivities to disruption, and the capabilities for identifying critical network components under various congestion levels.

To enable the optimization of transportation system resilience, quantifiable measures are needed. Zhang et al. (2009) proposed a Measure of Resilience quantification framework that combines functions of mobility, accessibility and reliability for intermodal transportation systems affected by major disruptions. That framework was extended by Zhang et al. (2010), in which effects of rerouting during performance recovery are simulated for evaluating resilience in integrated passenger and freight transportation systems. Miller-Hooks et al. (2012) defined network resilience as the expected fraction of demand that can be served after a disaster. They maximized it in a freight transportation network by employing two-stage stochastic integer formulations, in which preparedness decisions are made in the first stage and recovery actions for each disaster scenario are developed in the second stage. Later, using the same resilience indicator, Chen and Miller-Hooks (2012) quantified the inherent ability of

an intermodal freight network to recover from disruptions based on the network's topological and operational properties. They proposed a stochastic mixed-integer program for quantifying network resilience based on optimized post-disruption actions. Based on this measure of network resilience, Jin et al. (2014) developed a two-stage stochastic model, where the integration of bus services with the metro system was optimized to enhance the metro network resilience to potential disruptions. This resilience measure was also extended and applied to a container transportation network by Chen et al. (2017) from the perspective of network users (container shippers).

In a bi-level, three-stage stochastic programming model proposed by Faturechi and Miller-Hooks (2014), the resilience of a road network was defined as the post-response to pre-disruption ratio of the reciprocal of total travel time at a partial user equilibrium. The upper-level model optimized pre-disruption preparedness and post-disruption response decisions to maximize the expected resilience over possible scenarios. Ayyub (2014 and 2015) expressed network resilience with a network performance loss triangle over the period from the start of disruption to full recovery. Using this measure, a general framework was proposed by Zhang et al. (2018) for evaluating the resilience of large and complex rail transit networks by quantifying their vulnerability and recovery time with metrics and models based on network efficiency. Saadat et al. (2019) explored rail transit network resilience and its relations to vulnerability and efficiency, by using network topology analysis (i.e., of node connectivity) to determine network efficiency measures. The network efficiency was then used to derive vulnerability and resilience indexes. In analyzing the vulnerability of Washington DC's Metrorail system, they evaluated the effects of removing one component (i.e., station

or link) at a time and, unsurprisingly, the transfer stations were found to be the most vulnerable nodes. However, this study did not consider simultaneous failures of multiple components, budget constraints, or sequencing of the restoration actions.

Liu et al. (2020a) used three reliability measures based on connectivity, travel time, and capacity to evaluate the travel reliability of normal operations in a rail transit system. It was found that stations with higher connectivity reliability generally had higher capacity reliability, but lower travel time reliability. Liu et al. (2020b) explored the effects of transit line capacity reductions on system reliability, considering three indicators (i.e., crowding in trains, seat availability and perceived travel time). A measure defined as the Ratio of Affected Passenger Trips (RAPT) was proposed for identifying the critical links affecting the network's reliability. Both studies by Liu et al. evaluated the reliability of rail transit systems in relatively normal operations. An analysis of reliability reduction from major disruptions has not yet been found.

2.2 Post-disruption restoration

Restoration actions and operational changes after disruptions have been optimized for road and railroad networks. Karlaftis et al. (2007) used a genetic algorithm (GA) to allocate funds for urban road and bridge repairs after natural disasters. Kepaptsoglou et al. (2014) developed a model for optimizing the operations of a surviving road network after major disruptions by using shoulders and reversing some lanes. It used a bilevel approach in which the upper level selected the best combination of restoration actions using a GA, while the lower level used a traffic assignment model to determine traffic flow and evaluate system performance for each network configuration considered. Cacchiani et al. (2014) reviewed real-time models and algorithms for

managing and recovering from major disruption. Specifically, methods for train timetable rescheduling, rolling stock rescheduling, and crew scheduling, including their integration, were discussed.

For debris blockage events and similar disruptions in road networks, a common objective is to minimize the time needed to restore network connectivity, and the problem is usually formulated as a (mixed) integer programming model (Aksu and Ozdamar, 2014, Kasaei and Salman, 2016, Akbari et al., 2021a, Akbari et al., 2021b). Aksu and Ozdamar (2014) proposed a path-based dynamic model, where the objective is expressed as maximizing the sum of weighted earliness values of paths in each area (lower-level) and over all areas (upper-level), and limited restoration equipments are allocated to areas (upper-level). Kasaei and Salman (2016) formulated two models for two objectives: one for reconnecting the network in shortest time, and the other for maximizing benefit within a time limit. A variable neighborhood descent algorithm was developed for solution. Akbari et al. (2021a) optimized the post-disruption routing of multiple restoration work teams using a "Rich Local Search" heuristic, and Akbari et al. (2021b) explored the "online" optimization of restoration where the required time to restore a blocked link can only be known after in-site observations.

Post-disaster short-term network restoration activities are usually interrelated, and the sequence and schedule of their implementation needs to be optimized. Vugrin et al. (2014) constructed a bi-level model that involves network flows at the lower level and optimizes the sequence and schedule of post-disruption network recovery projects, where simulated annealing (SA) is used. The objective was to minimize the weighted sum of cumulative loss of system performance and total resource consumed during

recovery activities, considering task mode selection as well as precedence and resource constraints. Liu et al. (2021) developed a resilience-based model that optimizes restoration schedules of transportation network. A bi-objective model of network restoration scheduling subject to resource constraints was formulated, and heuristic algorithms were used to search non-dominated optimal restoration schedules. A travel speed-based metric was proposed to evaluate network performance and three independent indicators were developed to characterize network resilience. In a recent study for railroad networks, Wu et al. (2021) developed a model for optimizing the restoration sequence after major disruptions and solved it with a genetic algorithm. For that problem an exact solution through network flow reformulation was provided by Ng and Schonfeld (2023).

Post-disaster actions have also been studied for interdependent infrastructure networks. Sharkey et al. (2015) analyzed restoration interdependencies using the loss in restoration effectiveness resulting from individual, independent, and decentralized decision-making processes. They emphasized the importance of information sharing for interdependent infrastructure restoration to maximize the effectiveness of restoration of interdependent infrastructure networks. In a related network restoration study, Almoghathawi et al. (2019) proposed a multi-objective network restoration model for interdependent infrastructure networks, which sought to maximize the network resilience while minimizing the cost associated with the restoration process. Interdependence between power and water networks was analyzed. Sun and Zhang (2020) proposed a model, combining agent-based modeling and reinforcement

learning, for improving resilience of networks by exploring resource allocation (i.e., repair crew allocation strategy) for interdependent infrastructure networks.

2.3 Selection, sequencing, and scheduling of interrelated projects

The optimization problem of selecting, sequencing and scheduling interrelated projects dates back to the seminal work of Nemhauser and Ullmann (1969). In this early study, selection and scheduling of interrelated project investment was optimized under budget constraints, but the interrelation among projects was limited to pre-determined pairwise interaction coefficients which did not change over time. Later, Erlenkotter (1973) applied this problem to a hydroelectric system, where sequencing and scheduling of projects was optimized to minimize the present value (PV) of investment costs while meeting power requirements. This study considered project interrelation with upstream and downstream effects in a network instead of pairwise interaction, and these effects were also time-varying with increasing demands. However, there was no consideration of project durations. The sequencing of projects was in a fixed project set instead of a selected subset of candidate projects. The objective function did not include measures of network performance, nor did it consider a finite planning horizon. The interrelation of projects, simply defined by additional capacity parameters, was one-directional in the network.

From then on, there has been a vast literature exploring this field.

In a context of general projects, examples of such studies are as follows. Bouleimen and Lecocq (2003) explored the project scheduling problem and its multiple-mode variant that minimize total makespan under resource and precedence constraints. They proposed a customized simulated annealing (SA) algorithm that specifically considers

the solution space of project scheduling problems, with algorithm parameters based on results of statistical experiments on specific numeric examples. The objective function they use was a total duration of construction time instead of any economy-based value within a finite planning horizon. Mika et al. (2005) optimized the resource-constrained scheduling of multiple activities (similar to projects) with multiple modes for a maximized net present value (NPV) of cash flows in the planning horizon, where activities were represented as nodes in a network. For each activity only one mode must be selected. They considered four different payment models for NPV calculation, and compared performances of two heuristics they adopted for this problem: SA and tabu search. In the two studies above, interrelations among projects were only captured by precedence constraints. Tofighian and Naderi (2015) developed a novel mixed integer LP model for optimization of selection and scheduling of interrelated projects that jointly maximizes total time-varying expected benefit and minimizes variation of resource usage over the analysis period. A customized Pareto ant colony algorithm was proposed to solve this problem in discrete time steps. The project interrelation was captured by exclusivity constraints in specific subsets of projects. All three example studies above mainly intended to showcase the effectiveness of proposed solution methods. Both the computation of objective functions and the interdependency of projects were based on pre-determined parameters, and none of them involved changes in performance measures of underlying systems due to project implementations.

In various transportation networks, optimization of selection and scheduling of interrelated projects is also a crucial issue for planners to maximize total benefit or

minimize total cost over a specified planning horizon. Earlier applications of this problem in transportation include the following studies:

1) For inland waterways: Jong and Schonfeld (2001), Wang and Schonfeld (2005 and 2008).
2) For road (highway) networks: Hu and Schonfeld (1984), Wei and Schonfeld (1994), Tao and Schonfeld (2006 and 2007).

In more recent years, such studies include:

1) For road (highway) networks: Li et al. (2013), Bagloee and Asadi (2015), Miandoabchi et al. (2015), Gong and Fan (2016), Kumar and Mishra (2017), Hosseininasab et al. (2018), Jovanovic et al. (2018), Shayanfar and Schonfeld (2018 and 2019), Miralinaghi et al. (2020a), Miralinaghi et al. (2020b).
2) For rail (freight and transit) networks: Dao et al. (2019), Peng et al. (2019), Mohammadi et al. (2020).

The relevant studies on this problem are listed in Table 1, which classifies some characteristics of previous studies in this field. First, solution methods for this type of problem have been developed for several kinds of transportation networks, mostly using heuristic methods such as variations of genetic algorithms. Second, in more recent studies the problem formulation was usually bi-level, where flows in the network were assigned at the lower level and the selection and scheduling of projects was optimized at the upper level for a specific objective function (OF) computed over the planning horizon. In these bi-level models, the OFs were associated with costs, benefits, or travel time, and the project interrelation was captured by the projects' resulting changes in the flow-related network performances. In other words, the project

interrelation could not be explicitly expressed in OFs or constraints. Third, the implementation of projects was most commonly subject to budget constraints and was also affected by resource and precedence limitations. In most studies, however, the budget supply was limited to external sources. In many actual cases, funds can be collected from the network flows as an "internal" budget supply for project implementation. Fourth, the problem formulation mostly allowed budget and demand levels to be time-varying over the planning horizon, which was more commonly discrete than continuous. A discrete-time planning horizon facilitates computation of OFs, but limits the flexibility of determining project implementation times. Last, the problem was mostly deterministic, and only a few studies considered uncertain parameters. The performances of project implementation plans optimized for deterministic numerical cases are more vulnerable to unexpected future changes in parameters than those of robustly optimized plans considering multiple possible future scenarios. It should also be noted that in reality the uncertainties regarding multiple parameters concerning demand, resources, and construction costs are correlated. The existing studies either considers only one uncertainty or independent uncertainties.

Studies on this problem in the most recent years are highlighted as follows. Shayanfar and Schonfeld (2018) developed a method for selecting, sequencing and scheduling interrelated road projects in urban road networks. Interrelations among alternatives were captured by evaluating complete network models, while the project selection, sequencing and scheduling was optimized with a Genetic Algorithm (GA). That method was extended in Jovanovic et al. (2018) by scheduling improvements at intersections as well as links in urban street networks, and in Shayanfar and Schonfeld

(2019), by incorporating demand uncertainty in the formulation and optimizing the characteristics of network components being improved. The complete network was evaluated with a cell transmission model. A similar approach for selecting, sequencing and scheduling network improvements was applied in Peng et al. (2019) to rail transit network extensions. In both studies in 2019, the following features were considered: a continuous planning horizon, an internal supply of budget, a time-varying (growing) demand, and budget constraints that uniquely determined a schedule of project completion from a sequence. However, none of these studies explicitly considered durations of projects, let alone their possible overlaps. In practice, a project in progress could temporarily reduce capacities of affected network components, and the concurrent construction of multiple projects increases the possible network configurations (states) during the planning horizon. Correlated uncertainties were also neglected.

Dao et al. (2019) proposed a strategy for scheduling infrastructure renewal for a railway network, that minimized the sum of renewal cost and unavailability cost over a specified period. Instead of a heuristic search method, the strategy was a constrained three-step prioritization rule based on criticality of three levels, namely location, component type, and component. The prioritization-based solution, which was not statistically tested, had an unclear optimization quality. The renewal projects, without specified durations in each period, were recurrent for each network component, and their implementations were governed by possession (closure) time and due-date constraints. In a similar context, Mohammadi et al. (2020) developed a data-driven uncertainty set approximation for handling the uncertainty of recovery effects of

maintenance activities. A robust optimization of the selection and scheduling of renewal tasks was conducted, using commercial solvers with a greedy heuristic to provide a good initial feasible solution rather than a random start. Although the maintenance effects on track quality involved multiple parameters affecting that quality, the uncertainty set was based on observed data rather than probabilistic distributions and should be treated as a single uncertainty. In addition to possession time and maintenance threshold constraints, the implementations of recurrent renewal projects were also subject to budget and resource constraints. In both studies, the same project could be repeated, and the OFs did not involve costs or benefits induced by network traffic flows.

Table 1 Relevant studies on selection, sequencing and scheduling interrelated projects

Authors and date	Objective functions	Continuous planning horizon?	Optimizing selection & schedule?	Interrelated projects?	Solution methods	Applications	Constraints	Time varying?	Uncer-tainties?
Nemhauser & Ullmann (1969)	Max. present value (PV) of net worth (benefit) of projects	No	Yes	Yes	Dynamic programming	General investment	Budget	Budget	None
Erlenkotter (1973)	Min. PV of investment cost while meeting power requirements	Yes (infinite horizon)	No (sched. only)	Yes	Dynamic programming	Hydroelectric projects	Projected power requirements	Demand	None
Hu & Schonfeld (1984)	Improve level of service (traffic flow) by max. annual net benefit	No	No (sel. only)	Yes	IMSL routine ZXMIN	Various road projects & their combinations	None	Demand	None
Wei & Schonfeld (1994)	Min. PV of total user travel time costs plus project costs	No	Yes	Yes	Branch and bound, with artificial neural network for UE	Link capacity expansion (road)	Budget, project continuity	Demand	None
Jong & Schonfeld (2001)	Min. PV of total waiting time cost	Yes	Yes	Yes	GA	Lock improvement (waterway)	Budget	Budget, demand	None
Wang & Schonfeld (2005)	Max. PV of total user benefits	Yes	Yes	Yes	GA with a waterway simulation	Lock improvement (waterway)	Budget	Budget, demand	None
Tao & Schonfeld (2007)	Min. total user travel time (with a random term), based on UE traffic	No	Yes	Yes	Island model (GA extension)	Link improvement (road)	Budget	Budget, demand	Total travel time
Li et al. (2013)	Max. total benefit measured by net reductions of agency and user costs, based on traffic assigned by MMCN	No	No (sel. only)	Yes	Multi-commodity min. cost network, life-cycle cost ana-lysis, hypergraph Knapsack	Various road projects (widen-ing, new link, interchange…)	Budget	Budget, project cost, demand	Agency and user costs, dem. growth and discount rates
Bagloee & Asadi (2015)	Max. total benefit measured by total travel time reduced based on static traffic assignment	Yes	No (sched. only, se-quencing)	Yes	Gradient-based hybrid GA and ant colony on NN	Road addition	None	Demand	None
Miandoabchi et al. (2015)	Min. total travel time and CO emission (bi-objective), based on UE traffic	No	Yes	Yes	NSGA-II (a GA variant) and B-cell algorithm	Lane and road addition, lane allocation and alteration	Budget	Demand	None

Tofighian & Naderi (2015)	Max. total expected benefit and min. resource usage variation (bi-objective)	No	Yes	Yes	Customized ant colony optimization for MILP	General projects	Available resource, project inter-dependency	Available resource, project benefit	None
Gong & Fan (2016)	Min. excessive cost, based on UE traffic	No	No (sched. only)	Yes	GA, with F-W for UE	Mtn. (road)	A set of required projects	None	None
Kumar & Mishra (2017)	First optimize # of new lanes to min. resulting total travel cost, then jointly max. social benefit and min. total social cost	No	Yes	No	Force analogy; SPSA (Kumar & Peeta, 2014) for UE	Lane addition (road)	Budget	Budget	None
Hosseininasab et al. (2018)	Multi-objective: min. total travel time, max. user satisfaction over time, and max. spatial equity, based on UE traffic	No	Yes	Yes	2 approaches combining FW, GA, simplex phase I, & knees identification	Link addition (road)	Budget, technical limitation	Budget, demand	None
Dao et al. (2019)	Min. total renewal and unavailability costs, with economy of scale	No	Yes (recurrent)	Yes	Triple-prioritization rule	Infrastructure renewal in rail network	Possession time, network constraints, due-date	Maximum number of possession at location	None
Peng et al. (2019)	Min. PV of user cost plus supplier cost	Yes	Yes	Yes	Customized GA	Investments in rail transit network	Budget (with internal supply)	Demand	None
Shayanfar & Schonfeld (2019)	Min. PV of total system cost (including vehicle operation & safety)	Yes	Yes	Yes	GA, with F-W for assessment	Lane addition & widening (road)	Budget with fuel tax supply	Budget, demand	Demand growth
Miralinaghi et al. (2020a)	Min. total travel delay, based on UE traffic	No	Yes	Yes	"Active-set" algorithm	Road capacity improvement, link addition, mtn.	Budget	Budget, demand	None
Miralinaghi et al. (2020b)	Min. weighted total travel cost minus total business revenue, based on UE traffic	No	No (sched. only)	Yes	"Active-set" algorithm	Road capacity improvement, mtn.	Budget, required set of projects	Budget, demand	None
Mohammadi et al. (2020)	Max. weighted total quality of the network based on Track Quality Index	No	Yes (recurrent)	Yes	Greedy heuristic for an initial solu-tion to MILP and its robust version	Various rail freight network mtn.	Budget, mtn. thresholds, time allowed, resources	Constraint parameters, effects and costs of mtn.	Mtn. effects on track quality

mtn. = maintenance; max. = maximize; min. = minimize; sel. = selection; sched. = scheduling

2.4 Current research gap and presented contributions

The current research gap concerning transportation system resilience, transportation network restoration, and selection and scheduling of interrelated projects can be summarized as follows. First, most existing studies on network restoration, resilience, and vulnerability did not quantify the value of optimizing the sequencing and scheduling of restoration actions. A model for restoration sequence optimization is specifically missing for railroad freight networks. Second, existing studies on network restoration, including the ones that optimized the sequencing and scheduling of restoration actions, dealt with multiple deterministic disruption scenarios in a short-term static network configuration. When assessing long-term effects of network development, demand growth, probabilistic occurrence of disruption scenarios, and decisions on short-term network restoration should also be considered. Third, when optimizing the selection, sequencing and scheduling of interrelated projects for long-term network development, most existing studies considered deterministic parameters for the future. A few studies did consider some uncertain inputs, but none of these involved possible correlation among multiple uncertainties.

The presented contributions of the proposed dissertation work to the optimization of selection and scheduling of interrelated network projects are in the following aspects:

1) Optimization of short-term restoration sequence and schedule and long-term improvement plan are integrated into a tri-level model for a rail freight network. Using optimized short-term results and a measure of expected network resilience in the long-term problem, selection and scheduling of network improvement projects are also optimized.

2) In the rail freight network, the movement of multiple work teams in short-term restoration actions and the use of truck as a complementary shipment mode are considered.
3) In the road network, the proposed work is the first to combine the following features. Three correlated uncertain parameters are considered through the planning horizon for the long-term selection, sequencing, and scheduling (improvement plan) of projects. A number of scenarios are generated based on the distribution of these parameters, and the evaluation of candidate long-term improvement plans is based on the expected objective function value over these scenarios. Internal as well as external budget sources and a continuous planning horizon are considered jointly, allowing more flexibility in deciding when to start and complete the projects. The demand growth and its uncertainty are considered for both road and rail networks.
4) In the road network, the use of two modes of traffic – bus and car – is allowed in the model. The iteration of mode shares is integrated with the lower-level traffic assignment, influencing upper-level optimization.
5) In both problems, implementation periods of projects can overlap. In the road network problem, an improvement project in progress reduces the capacities of affected links. Over-lapping project implementations result in more possible network configuration states.

Chapter 3: Problem formulation

3.1 Problem 1: Integrated model for rail freight network

The notations for the variables used in Problem 1 are shown in Table 2.

Table 2 Notations for variables in Problem 1

Symbol	Description	Unit	Baseline value
c_{alt}	Unit shipment cost by an alternative mode	\$/ton/mi	2.5
c_{alt}^{w}	Unit shipment cost by an alternative mode for OD pair w	\$/ton	
c_r	Unit rail shipment cost	\$/ton/mi	0.5
c_u	Unit value of users' shipment time	\$/ton/hr	10
C_l	Total construction cost of improvement project l	\$	
d_a	Length of link a	mile	
E_{κ_i}	Monthly expected excess under network configuration κ_i	\$/month	
F	Monthly external budget supply	\$/month	5×10^6
$k_{a(n)}^{0}$	Normal capacity of link a (node n)	ton/hr	
$k_{a(n)}^{s}$	Capacity of link a (node n) in stage s during restoration	ton/hr	
$k_{a(n),\omega}$	Capacity of link a (node n) when damaged in scenario ω	ton/hr	
q_w	Amount of freight that must be delivered for OD pair w	ton/hr	
r	Monthly interest rate		0
t_a^0	Normal travel time through link a	hour	
t_a^s	Travel time through link a in stage s during restoration	hour	
$t_{a,\omega}$	Travel time through link a when damaged in scenario ω	hour	
T	Duration of the planning horizon	month	300
T_w	Maximum rail shipment time allowed for OD pair w	hour	
X_{ω,κ_i}	Cumulative cost increment (excess) in the restoration after disruption scenario ω under network configuration κ_i	\$	
Z_s	Total hourly cost in stage s during restoration	\$/hr	
$\gamma_{a(n),p}^{w}$	Binary variable indicating whether link a (node n) is used by path p of OD pair w		
δ_s	Duration of stage s in a restoration process	hour	
$\theta_p^{w,s}$	Fraction of demand for OD pair w to be shipped through path p during stage s		
π_ω	Monthly occurrence probability of disruption scenario ω	1/month	
τ	Time elapsed since the start of the planning horizon	month	
τ_i	Starting time of network configuration κ_i	month	
φ_i	Duration of network configuration κ_i	month	
ϕ_l	Required (minimum) work duration of improvement project l	month	

3.1.1 Lower and middle-level model for optimizing short-term restoration sequence

In a rail freight network, there are nodes $n \in N$ and rail links $a \in A$. For each OD pair $w \in W$ in this network, q_w tons of freight must be delivered per hour through multiple connecting paths $p \in P_w^s$.

The set of disruption scenarios is denoted as Ω. In a disruption scenario $\omega \in \Omega$, multiple network components are damaged, with damaged links in set A_ω and nodes in set N_ω. The number of disrupted components is denoted as $|S_\omega|$. Multiple possible restoration sequences exist, with limits on labor, budget and other resources. It is assumed here that a "restoration sequence" is specified by a one-dimensional permutation of damaged components that is later used for uniquely determining the restoration schedule for multiple work teams. Under this assumption, there are $|S_\omega|!$ possible restoration sequences given the disruption scenario ω. For each sequence $\sigma_{res} \in \Sigma_{re\ ,\omega}$, the whole restoration process is divided into $|S_\omega|$ stages. Each stage $s \in S_\omega$ has its duration δ_s, which corresponds to the time period between the $(s-1)^{\text{th}}$ and the s^{th} completion of recovering a component (or between the onset of disruption and the recovery of the first component). In addition, "stage 0" in the set S_ω represents normal operation of the network.

The capacity of each network component is defined as the largest possible hourly amount of freight that can be shipped through that component. Each link (node) has its capacity k_a^s (k_n^s) in stage s. k_a^s (k_n^s) equals to k_a^0 (k_n^0) under normal condition and decreases to $k_{a,\omega}$ ($k_{n,\omega}$) if it is damaged in scenario ω. Travel time through each link in stage s, denoted as t_a^s, equals to t_a^0 in normality and increases to $t_{a,\omega}$ under disruption. The capacity and travel time of each disrupted component return to normal

after restoration actions are completed for that component. For each stage s and each OD pair w, simple paths with total travel time shorter than T_w are in the path set P_w^s. For OD pair w, if there is no remaining capacity for any path in P_w^s, the unserved demand is assumed to be immediately shipped through an alternative mode (e.g., trucks). Besides, each link has its length d_a.

For a given disruption scenario ω, total hourly cost Z_s is minimized for each stage s in S_ω (including stage 0, the normal state). Before full recovery of the network, reduced capacity and increased travel time on damaged components result in the excess of minimized total hourly cost (Z_s, $s \neq 0$) over that in the normal state (Z_0). Through the whole recovery process this excess yields a cumulative increment of cost, which is defined as "excess" and is denoted as $X_\omega(\sigma_{res})$, a function of σ_{res}. The middle-level model (containing the lower level) aims at finding the excess-minimizing restoration sequence σ_{res}, whose resulting excess measures short-term resilience of this network under disruption scenario ω. The model is formulated as follows:

$$Min\ X_\omega(\sigma_{res}) = \sum_{s \in S_\omega / \{0\}} \delta_s (Z_s - Z_0)\,, \sigma_{res} \in \Sigma_{res,\omega} \tag{1}$$

subject to:

$$Z_s = Min \sum_{w \in W} [\sum_{p \in P_w^s} \sum_{a \in p} \gamma_{a,p}^w \theta_p^{w,s} q_w\,(c_r t_a^s + c_u d_a) + c_{alt}^w (1 - \sum\nolimits_{p \in P_w} \theta_p^{w,s}) q_w]\,, \forall s \in S_\omega \tag{2}$$

$$\sum\nolimits_{p \in P_w^s} \theta_p^{w,s} \leq 1,\ \ \forall s \in S_\omega, \forall w \in W \tag{3}$$

$$\sum_{w \in W} \sum_{p \in P_w^s} \gamma_{a(n),p}^w\, \theta_p^{w,s} q_w \leq k_{a(n)}^s,\ \forall s \in S_\omega,\ \forall a \in A (n \in N) \tag{4}$$

$$\theta_p^{w,s} \in [0,1],\ \forall s \in S_\omega, \forall\, p \in P_w^s,\ \forall w \in W \tag{5}$$

Decision variable $\theta_p^{w,s}$ is the fraction of demand for OD pair w to be shipped through path p during stage s. Binary indicator $\gamma_{a,p}^w$ ($\gamma_{n,p}^w$) equals 1 if link a (node n) is used by path p. Equation (2) defines Z_s as the minimized hourly total cost in stage s, which includes the supplier's rail shipment cost, the users' value of shipment time, and the supplier's alternative mode (e.g., truck) shipment cost, whose unit values are denoted as c_r, c_u, and c_{alt}^w, respectively. For each OD pair with demand, constraint (3) ensures that the total fraction of demanded freight shipped by rail does not exceed 1. Constraint (4) ensures that total freight flow through each network component does not exceed its capacity. Constraint (5) determines domains for decision variables. (3), (4), and (5) serve as constraints for minimizing Z_s in (2) at the lower level.

3.1.2 Upper-level model for selecting and scheduling long-term improvements

In a planning horizon of T months, there are $|\Lambda|$ available projects for improving a rail freight network, whose set is denoted as Λ. An improvement project may increase capacities of existing network components or add new links in the network. The number of projects L to-be-implemented in an improvement sequence satisfies $0 \leq L \leq |\Lambda|$. Then, there are $|\Lambda|!/(|\Lambda| - L)!$ possible sequences with L improvements. The set of all possible improvement sequences is denoted as Σ_{imp}, and the selected sequence is denoted as $\sigma_{imp} \in \Sigma_{imp}$, with $\left|\sigma_{imp}\right| = L$ improvements in the sequence. The i^{th} improvement in the selected sequence is denoted as σ_{imp}^i, if $\left|\sigma_{imp}\right| \neq 0$. Completion of each improvement leads to changes in network configuration. The original network configuration is denoted as κ_0, and updated configurations following the first, second, and later completions of improvements as $\kappa_1, \kappa_2, \ldots, \kappa_{|\sigma_{imp}|}$. These configurations last

for $\varphi_0, \varphi_1, \varphi_2, \ldots, \varphi_{|\sigma_{imp}|}$ months, respectively. With τ denoting the number of months elapsed since the start of the planning horizon, the starting times of these configurations are $\tau_0, \tau_1, \tau_2, \ldots, \tau_{|\sigma_{imp}|}$, respectively.

Using the middle-level model, an optimized sequence of short-term restoration and its resulting excess (denoted as X_{ω,κ_i}) are found for each disruption scenario ω under network configuration κ_i. Since the network configuration in each phase depends on the selected improvement sequence σ_{imp}, the excess X_{ω,κ_i} is further denoted as a function of σ_{imp}, which is $X_{\omega,\kappa_i}(\sigma_{imp})$.

In the optimization of long-term improvements, the following assumptions are applied:

1) Throughout the planning horizon, the set of possible disruption scenarios is fixed. The demand for all OD pairs grows exponentially at a constant annual rate (including zero).
2) The monthly occurrence probability π_ω of a disruption scenario ω is small enough (below 5%) that concurrence of multiple scenarios and recurrence of a certain scenario are negligible.
3) Each improvement project $l \in \Lambda$ has its total implementation (construction) cost C_l, and the cost is paid in full upon completion.
4) The initial available budget is zero. During the planning horizon, there is a constant monthly external supply of budget F, and the available budget cannot be negative at any time.

For the upper-level model, to optimize the selection and schedule of improvements based on network resilience, the sequence σ_{imp} that minimizes the expected excess

plus total improvement cost incurred in the planning horizon must be found. This can be formulated as:

$$Min\ TC(\sigma_{imp} \in \Sigma_{imp}) = \begin{cases} \sum_{i=1}^{|\sigma_{imp}|} C_{\sigma_{imp}^i} + \sum_{i=0}^{|\sigma_{imp}|} \varphi_i E_{\kappa_i}(\sigma_{imp}), & \text{if } |\sigma_{imp}| \neq 0 \\ TE_{\kappa_0}, & \text{if } |\sigma_{imp}| = 0 \end{cases} \quad (6)$$

subject to:

$$E_{\kappa_i}(\sigma_{imp}) = \sum_{\omega} \pi_\omega X_{\omega,\kappa_i}(\sigma_{imp}), \qquad \forall 0 \leq i \leq |\sigma_{imp}| \quad (7)$$

$$\sum_{i=1}^{|\sigma_{imp}|} C_{\sigma_{imp}^i}/F < T, \qquad \text{if } |\sigma_{imp}| \neq 0 \quad (8)$$

Equation (7) specifies the monthly expected excess under the network configuration κ_i in each phase. Constraint (8) filters out sequences whose last improvement cannot be completed within the planning horizon.

If the present value (PV) is considered in the objective function (6), with a monthly interest rate of r, the modified objective function is:

$$Min\ TC(\sigma_{imp} \in \Sigma_{imp})$$

$$= \begin{cases} \sum_{i=1}^{|\sigma_{imp}|} \frac{C_{\sigma_{imp}^i}}{(1+r)^{\tau_i}} + \sum_{i=0}^{|\sigma_{imp}|} \int_{\tau_i}^{\tau_i+\varphi_i} \frac{E_{\kappa_i}(\sigma_{imp})}{(1+r)^{\tau}} d\tau, & \text{if } |\sigma_{imp}| \neq 0 \\ \int_0^T \frac{E_{\kappa_0}}{(1+r)^{\tau}} d\tau, & \text{if } |\sigma_{imp}| = 0 \end{cases} \quad (6')$$

3.1.3 Rules for determining the schedule with a given sequence of restoration/improvement

First, the following set of rules translates a given short-term restoration sequence σ_{res} into a unique schedule by applying the binding constraint on available work teams. By these rules, each damaged component is restored as soon as possible so that the excess is minimized for a given restoration sequence and a disruption scenario.

To restore damaged components in response to the disruption event, multiple work teams (may also be only one team) are available. The following assumptions are made in analyzing restoration actions:

1) Once a disruption occurs, all work teams depart from a single base node. After complete restoration of all damaged components, they return to the same node.
2) The required time to fully restore a certain damaged component is the same for all work teams.
3) Regardless of the disrupted status of components, all work teams move in the network at the same speed.
4) For determining access times from and to a link, each link is represented by its midpoint (thus treating a link as a node) as a simplification. The shortest path is found from the origin "node" to the destination "node" in the railroad network.

With each work team restoring one component at a time, the largest number of concurrently restored components equals the number of work teams. When determining completion times of restoration stages with a given sequence, the access time for work teams to move from one network component to another is considered. To ensure that a restoration sequence uniquely determines the schedule for work teams, the following set of rules are used. First, with work teams labeled 1, 2, 3, …, the 1st, 2nd, 3rd, … components in a sequence are assigned to teams 1, 2, 3, …, respectively. Then, as soon as the restoration of a component is completed by a certain team, this team proceeds to the first component in the sequence that has no ongoing or completed restoration. Under this rule the duration of each stage (δ_s) is determined for the given

sequence. For example, nodes 2, 4, 3, and 1 need to be restored sequentially, and work teams 1 and 2 are available. Team 1 is assigned Node 2 as its first restoration task, while Team 2 must deal with Node 4. Restoration of Node 2 is completed with that of Node 4 still in progress, and Team 1 starts restoring Node 3 after accessing Node 3 from Node 2. Before restoration of Node 3 ends, Team 2 finishes its task at Node 4 and heads for Node 1. Restoration of Node 1 is completed earlier than that of Node 3. In this example δ_1 is the duration from the onset of disruption to the end of Node 2 restoration, while δ_2, δ_3, δ_4 are given by time periods between completions of restoring nodes 2 and 4, 4 and 1, 1 and 3, respectively.

For uniquely determining the schedule of improvement projects with a given improvement sequence σ_{imp}, the following set of rules is applied based on the binding constraints on budget and required construction times.

It should be noted that multiple projects can proceed concurrently. Let σ_{imp}^{i} be the i^{th} project in this sequence to start in the schedule. Each project has its construction cost $C_{\sigma_{imp}^{i}}$ as well as its required (minimum) work duration $\phi_{\sigma_{imp}^{i}}$. With a constant external monthly budget supply F, it takes $\max\{C_{\sigma_{imp}^{i}}/F,\ \phi_{\sigma_{imp}^{i}}\}$ months to complete construction of σ_{imp}^{i}, which means that a project cannot be completed until its funding is fully ready. Whenever funding is completed for project σ_{imp}^{i} (which is $C_{\sigma_{imp}^{i}}/F$ months after its commencement), the next project in the sequence (σ_{imp}^{i+1}, if exists) starts. If completion of a project decreases the monthly expected excess, this project should be started and completed as soon as possible to minimize the cumulative expected excess in the planning horizon. If work duration is not yet reached for σ_{imp}^{i},

concurrent construction of σ_{imp}^{i} and σ_{imp}^{i+1} occurs. The duration of phases (φ_i) and their corresponding network configurations (κ_i) depend on completion times of improvement projects. It is defined that Phase i starts from the i^{th} completion time of all projects in the sequence σ_{imp} and ends at the $(i+1)^{\text{th}}$ completion time. Phase 0 lasts from the start of planning horizon to the first completion time of all projects, while Phase $|\sigma_{imp}|$ stretches from the final completion of all projects to the end of planning horizon. An example of determining the schedule for a sequence with 5 selected projects is shown in Figure 1, with given input parameters $C_{\sigma_{imp}^{i}}$, F, and $\phi_{\sigma_{imp}^{i}}$. For the second and fourth project in the sequence, $C_{\sigma_{imp}^{i}}/F \geq \phi_{\sigma_{imp}^{i}}$, while for other projects $C_{\sigma_{imp}^{i}}/F < \phi_{\sigma_{imp}^{i}}$.

3.2 Problem 2: Optimizing road network improvement

The notations for the variables used in Problem 2 are shown in Table 3.

Table 3 Notations for variables in Problem 2

Symbol	Description	Unit	Baseline value
b	Equivalent number of cars per bus in traffic flow		2
c_{bus}	Bus fare	\$/person	1.2
c_{fuel}	Unit car fuel cost	\$/veh/mi	0.06
C_l	Total construction cost of improvement project l	\$	
d_a	Length of link a	mile	
d_p	Length of path p	milt	
f_{bus,R_i}	Required hourly bus service frequency for bus route R_i	veh/hr	
f_{car}^{w}	Total hourly number of cars for OD pair w	veh/hr	
f_p^w	Hourly number of vehicles using path p of OD pair w	veh/hr	
F	Annual external budget supply	\$/year	1.5×10^7
g	Annual growth rate of demand		2.5%
h_{bus,R_i}	Bus headway for bus route R_i	hour	
H	Number of effective hours in a year		8,760

I_w^{bus}	Impedance for each bus passenger of OD pair w	\$/person	
I_w^{car}	Impedance for each car user of OD pair w	\$/person	
$I_{w,p}^{car}$	Impedance for each car user using path p of OD pair w	\$/person	
k_a	Actual capacity of link a	veh/hr	
k_{bus,R_i}	Capacity of dedicated bus lane for bus route R_i	veh/hr	
K_{bus}	Passenger capacity of a bus		15
K_{car}	Average number of people in a car		1.25
$q_{w,P(OP)}$	Number of vehicles of OD pair w per hour during peak (off-peak) hours	veh/hr	
r	Annual interest rate		5%
t_a	Travel time through link a	hour	
t_a^0	Free-flow travel time through link a	hour	
t_{wait,R_i}	Average wait time of a passenger using bus route R_i	hour	
T	Duration of the planning horizon	year	15
v	Unit value of vehicles' travel time (Unit value of users' travel time when bus is allowed)	\$/veh/hr (\$/psn/hr)	15(12)
v_{wait}	Unit value of bus passengers' wait time	\$/psn/hr	18
x_a	Hourly traffic flow (number of vehicles) through link a	veh/hr	
$Y_{\omega,\sigma_{imp}}$	Present value of cumulative total cost of travel time over the planning horizon with the improvement sequence σ_{imp} in scenario ω	\$	
$Z_{P(OP)}$	Hourly total travel time cost in the network during peak (off-peak) hours	\$/hr	
Z	Average hourly total travel time cost	\$/hr	
α	Fraction of peak hours in all hours		0.25
β_l^{ω}	Binary variable indicating whether project l can be completed within the planning horizon in scenario ω		
$\gamma_{a,p}^{w}$	Binary variable indicating whether link a is used by path p of OD pair w		
$\zeta_{w,P(OP)}^{bus}$	Share of bus passengers in all travelers of OD pair w during peak (off-peak) hours		
$\zeta_{w,P(OP)}^{car}$	Share of car users in all travelers of OD pair w during peak (off-peak) hours		
$\boldsymbol{\mu}$	Mean vector of uncertain variables		
μ_i	Mean of variable i in the distribution $\mathcal{F}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$		
ρ	Fraction of total travel time cost used as an internal source of budget		0.01
ρ_{ij}	Correlation coefficient between variables i and j in distribution $\mathcal{F}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$		
$\boldsymbol{\Sigma}$	Correlation matrix of uncertain variables		
σ_i	Standard deviation of variable i in distribution $\mathcal{F}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$		
τ	Time elapsed since the start of the planning horizon	year	

τ_i^ω	Starting time of network configuration κ_i in scenario ω	year	
τ_l^ω	Budget-ready time of improvement project l in scenario ω	year	
φ_i^ω	Duration of network configuration κ_i in scenario ω	year	
ϕ_l	Required (minimum) work duration of improvement project l	year	
χ_p^w	Fraction of car users of OD pair w who use path p		
ψ	Multiplier of the required construction time of projects		1.0

3.2.1 Lower-level traffic assignment model

In a road network, there are nodes $n \in N$ and links $a \in A$. Each link has its length d_a in miles. For each OD pair $w \in W$ in this network, during peak hours $q_{w,P}$ vehicles/hour travel through multiple connecting paths $p \in P_w^s$. During off-peak hours, this number of vehicles is denoted as $q_{w,OP}$.

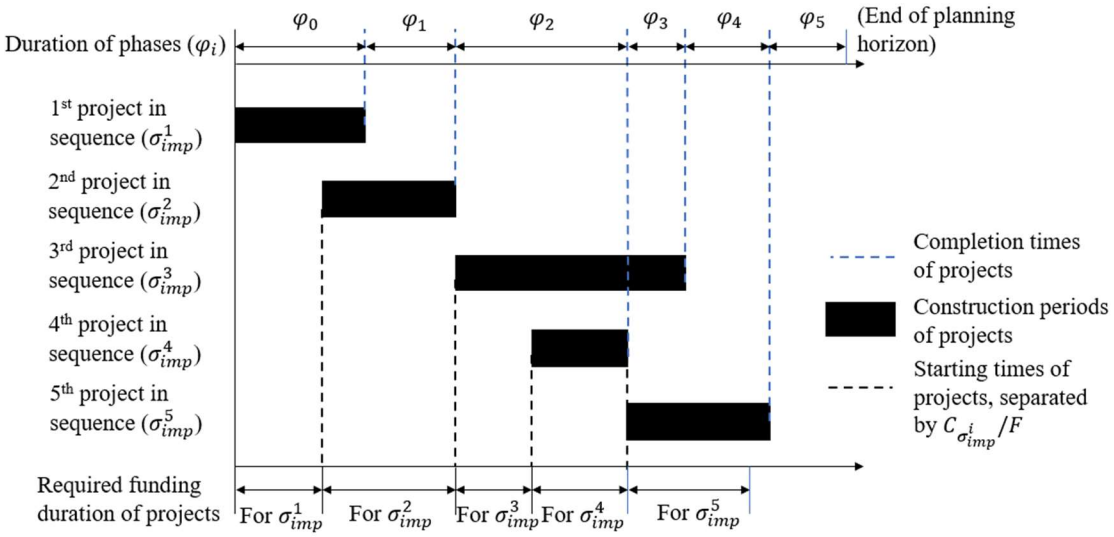


Figure 1 Determining the schedule for a sequence with 5 projects (in Problem 1)

Each link has its actual capacity k_a in vehicles per hour. Travel time through each link in hours, denoted as t_a, equals t_a^0 when there are no vehicles on the link. With a

traffic flow of x_a in vehicles per hour, the travel time t_a is given by the following congestion function proposed by the Bureau of Public Roads (BPR, 1964):

$$t_a(x_a) = t_a^0[1 + 0.15\frac{x_a}{k_a}]^4 \tag{9}$$

For each OD pair w, the set of all possible simple paths in the network is denoted as P_w, and the number of vehicles in an hour using path $p \in P_w$ is denoted as f_p^w. According to the user equilibrium (UE) conditions proposed by Wardrop (1952), the following nonlinear programming problem is formulated for the peak-hour (off-peak) traffic:

$$Min \sum_{a \in A} \int_0^{x_a} t_a(x)dx \tag{10}$$

subject to:

$$\sum_{p \in P_w} f_p^w = q_{w,P(OP)}, \forall w \in W \tag{11}$$

$$f_p^w \geq 0, \forall p \in P_w, w \in W \tag{12}$$

$$x_a = \sum_{w \in W} \sum_{p \in P_w} f_p^w \gamma_{a,p}^w, \forall a \in A \tag{13}$$

The objective function to be minimized in (10) is the sum of integrals of the link congestion functions with respect to traffic flows. Constraints (11) and (12) ensure that for each OD pair, the sum of all non-negative traffic flows on different paths equals the OD pair's total traffic demand. The traffic flow through each link is given by constraint (13), where $\gamma_{a,p}^w$ is a binary indicator that equals 1 only if link a is used by path $p \in P_w$ of OD pair w. When the optimized UE assignment is reached, no vehicle can reduce its travel time from its origin to its destination by shifting its route. The UE traffic flow

on link a during a peak (off-peak) hour is denoted by $x_{a,P(OP)}^{UE}$. The hourly total travel time cost in the peak (off-peak) network is given by:

$$Z_{P(OP)}^{UE} = v \sum_{a \in A} x_{a,P(OP)}^{UE} t_a\left(x_{a,P(OP)}^{UE}\right) \tag{14}$$

where v is the unit value of vehicles' time in dollars per vehicle hour. If the fraction of peak hours in all hours is $\alpha \in [0,1]$, then the average hourly total travel time cost is given by:

$$Z_{UE} = \alpha Z_P^{UE} + (1-\alpha) Z_{OP}^{UE} \tag{15}$$

3.2.2 Sampling scenarios with correlated uncertain variables

If a selected set of input parameters for the model are considered as correlated uncertain variables, they follow a certain type of continuous multivariate distribution as given by:

$$(param_1, param_2, param_3, \dots)^T \sim \mathcal{F}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{16}$$

where $\boldsymbol{\mu}$ is the mean vector, and $\boldsymbol{\Sigma}$ is the correlation matrix. All elements in these tensors are constant parameters. From the continuous multivariate distribution specified above, a preset number of vectors of uncertain variables are sampled. Each sampled vector represents a combination scenario $\omega \in \Omega$ of uncertain variables. The upper-level evaluation of selection and schedule of projects is based on the same set of generated scenarios. In the numerical case to be discussed in Section 6, three correlated uncertain variables are considered: annual growth rate of demand (g, assumed to be the same for all OD pairs), external annual budget supply (F), and the multiplier of the required construction time of projects (ψ).

3.2.3 Upper-level model for optimizing selection and scheduling of improvement projects

In a planning horizon of T years, there are $|\Lambda|$ available projects for improving the road network, whose set is denoted as Λ. These projects include increasing capacities of existing network components or adding new links in the network. The number of projects L to-be-implemented in an improvement sequence satisfies $0 \leq L \leq |\Lambda|$. Then, there are $|\Lambda|!/(|\Lambda| - L)!$ possible sequences with L improvements. The set of all possible improvement sequences is denoted as Σ_{imp}, and the selected sequence is denoted as $\sigma_{imp} \in \Sigma_{imp}$, with $\left|\sigma_{imp}\right| = L$ improvements in the sequence. The i^{th} improvement in the selected sequence is denoted as σ_{imp}^{i}, if $\left|\sigma_{imp}\right| \neq 0$. Considering that capacities of some links are reduced during construction of a project, both the start and the completion of each improvement lead to changes in network configuration (including availability of nodes and links as well as link capacities). The original network configuration is denoted as κ_0, and the updated configurations following the first, second, and later starts/completions of improvements are denoted as κ_1, κ_2, …, $\kappa_{2|\sigma_{imp}|}$. In scenario ω, these configurations start at τ_0^ω, τ_1^ω, τ_2^ω, …, $\tau_{2|\sigma_{imp}|}^\omega$ years into the planning horizon, and they last for φ_0^ω, φ_1^ω, φ_2^ω, …, $\varphi_{2|\sigma_{imp}|}^\omega$ years, respectively.

In a generated scenario ω with a sampled vector of uncertain variables $(g_\omega, F_\omega, \psi_\omega)^T$, the peak (off-peak) hourly demands of all OD pairs increase at a constant annual rate:

$$q_{w,P(OP)}(\tau) = (1 + g_\omega)^\tau q_{w,P(OP)}(0) \tag{17}$$

where τ is the number of years elapsed since the start of the planning horizon. $q_{w,P(OP)}(0)$ denotes the hourly demand of OD pair w at the year zero, and this is a

fixed parameter. Also in this scenario ω, the external annual budget supply is F_ω, and the original required (minimum) work duration ϕ_l of each project $l \in \Lambda$ is multiplied by ψ_ω.

With a time-varying demand level, the average hourly total cost of travel time under user equilibrium (Z_{UE}) as obtained at the lower-level model can be treated as a function of τ. For a given improvement sequence σ_{imp}, different scenarios ω lead to different time spans of network configuration κ_i ($0 \le i \le 2|\sigma_{imp}|$) as well as different growth rates of demand. As a result, the value of $Z_{UE}(\tau)$ is also subject to changes of κ_i, ω, and σ_{imp}. Let $Z_{UE}^{\omega,\kappa_i,\sigma_{imp}}(\tau)$ be a continuous function that is integrable in the time interval where the network is in phase i with configuration κ_i, given the evaluated improvement sequence σ_{imp} and the scenario ω. The present value (PV) of cumulative total cost of travel time over the planning horizon with the sequence σ_{imp} and the scenario ω is:

$$Y_{\omega,\sigma_{imp}} = H \sum_{i=1}^{2|\sigma_{imp}|} \int_{\tau_{i-1}^{\omega}}^{\tau_i^{\omega}} \frac{Z_{UE}^{\omega,\kappa_i,\sigma_{imp}}(\tau)}{(1+r)^{\tau}} d\tau \tag{18}$$

where H is the number of effective hours in a year, and r is the annual interest rate. This equation serves as a general formulation, while the actual computation of $Y_{\omega,\sigma_{imp}}$ uses the approximation method that will be provided in Section 4.1.

Given the evaluated sequence σ_{imp} of improvement projects, the total implementation (construction) cost of the i^{th} project in the sequence is denoted as $C_{\sigma_{imp}^i}$. Under the same improvement sequence, in scenario ω, whether a project can be completed within the planning horizon is subject to the values of F_ω and ψ_ω. Let $\beta_{\sigma_{imp}^i}^{\omega}$

be a binary indicator that is 1 if the i^{th} project in the sequence σ_{imp} can be completed within the planning horizon in scenario ω. The detailed rules for determining $\beta^{\omega}_{\sigma^i_{imp}}$ are explained in Section 3.2.4.

In the upper-level model, to optimize the selection and schedule of improvements, the objective is to find the sequence σ_{imp} that minimizes the expected PV of the sum of cumulative travel time cost plus total project implementation cost (i.e., PVC) during the planning horizon:

$$Min\ PVC(\sigma_{imp} \in \Sigma_{imp}) = |\Omega|^{-1} \sum\nolimits_{\omega \in \Omega} (Y_{\omega,\sigma_{imp}} + \sum\nolimits_{i=1}^{|\sigma_{imp}|} \beta^{\omega}_{\sigma^i_{imp}} \frac{C_{\sigma^i_{imp}}}{(1+r)^{\tau^{\omega}_{\sigma^i_{imp}}}}) \quad (19)$$

where $\tau^{\omega}_{\sigma^i_{imp}}$ denotes the time when the cumulative available budget becomes sufficient to pay for the i^{th} project in the sequence σ_{imp} under the scenario ω. This can be interpreted as finding the weighted average of total cost over all sampled scenarios that are equally weighted. It is assumed that the payment from the available budget for project σ^i_{imp} occurs only at the time $\tau^{\omega}_{\sigma^i_{imp}}$. Rules for determining $\tau^{\omega}_{\sigma^i_{imp}}$ are explained in Section 3.2.4. It should be noted that, for a given sequence σ_{imp}, its i^{th} project σ^i_{imp} is also the i^{th} project to start in the planning horizon, but not necessarily the i^{th} project to be completed.

3.2.4 Constraints and rules for scheduling project implementation

In this section, constraints and rules for uniquely determining starting and completion times of projects under a given scenario ω and a given improvement sequence σ_{imp} are presented. By these rules, for each project whose completion

decreases the average hourly travel time cost, it is started and completed as soon as its cost is justified by its benefit under the demand level and the binding constraints allow, so that the cumulative savings in travel time cost over the planning horizon is maximized. The following assumptions are applied:

1) Each improvement project $l \in \Lambda$ has its total implementation (construction) cost C_l, and the cost is paid in full as soon as its funding is fully ready.
2) The initial available budget is zero. During the planning horizon, there is a constant monthly external supply of budget F_ω, and available budget cannot be negative at any time.
3) A project cannot be completed until its funding is fully ready.
4) A constant small fraction ρ of total travel time cost is used as an internal source of budget.

Similarly to the rail freight network presented in Section 3.1, concurrent construction of multiple projects is allowed. The i^{th} project to start in the sequence σ_{im} , denoted as σ_{imp}^{i}, has its construction cost $C_{\sigma_{imp}^{i}}$ and its required work duration $\psi_\omega \phi_{\sigma_{imp}^{i}}$. If there is only external budget supply, then $\max\{C_{\sigma_{imp}^{i}}/F_\omega,\ \psi_\omega \phi_{\sigma_{imp}^{i}}\}$ years are needed to complete the construction of σ_{imp}^{i}. The times when available budget is ready for the i^{th} project in the sequence σ_{imp} are given by

$$\tau_{\sigma_{imp}^{i}}^{\omega} = C_{\sigma_{imp}^{i}}/F_\omega + \tau_{\sigma_{imp}^{i-1}}^{\omega}, \qquad 1 \le i \le \left|\sigma_{imp}\right| \tag{20a}$$

If internal budget supply is available, however, $\tau_{\sigma_{imp}^{i}}^{\omega}$ is given by

$$C_{\sigma_{imp}^{i}} = \sum_{j=j_1}^{j_2} \int_{\tau_j^{\omega}}^{\tau_{j+1}^{\omega}} [F_{\omega} + \rho HZ_{UE}^{\omega,\kappa_j,\sigma_{imp}}(\tau)] d\tau, \qquad 1 \le j \le 2\left|\sigma_{imp}\right| \tag{20b}$$

subject to:

$$\tau_{j_1-1}^{\omega} < \tau_{\sigma_{imp}^{i-1}}^{\omega} = \tau_{j_1}^{\omega}, \qquad 2 \le i \le \left|\sigma_{imp}\right| \tag{21a}$$

$$\tau_{j_2}^{\omega} < \tau_{\sigma_{imp}^{i}}^{\omega} = \tau_{j_2+1}^{\omega}, \qquad 1 \le i \le \left|\sigma_{imp}\right| \tag{21b}$$

$$\tau_{\sigma_{imp}^{0}}^{\omega} = 0 \tag{21c}$$

where $\tau_{\sigma_{imp}^{i}}^{\omega}$ needs to be numerically determined using the methods provided in Section 4. With internal and external budget supply, $\max\{\tau_{\sigma_{imp}^{i}}^{\omega} - \tau_{\sigma_{imp}^{i-1}}^{\omega}, \psi_{\omega}\phi_{\sigma_{imp}^{i}}\}$ years are needed to complete the construction of σ_{imp}^{i}.

Whenever funding is completed for project σ_{imp}^{i}, the next project σ_{imp}^{i+1} in the sequence (if it exists) starts. If work duration is not yet reached for σ_{imp}^{i}, concurrent construction of σ_{imp}^{i} and σ_{imp}^{i+1} occurs. The durations of phases (φ_i^{ω}) and their corresponding network configurations (κ_i) depend on starting and completion times of improvement projects. It is defined that Phase i starts from the i^{th} starting/completion time (τ_i^{ω}) of all projects in the sequence σ_{imp} and ends at the $(i+1)^{\text{th}}$ starting/completion time (τ_{i+1}^{ω}). If the starting/completion time of one project overlaps that of another project (that is, $\tau_{i_1}^{\omega} = \tau_{i_2}^{\omega}, i_1 < i_2$), then there is a corresponding phase with zero duration ($\varphi_{i_1}^{\omega} = 0$). If the improvement sequence is not empty (i.e., $\left|\sigma_{imp}\right| \neq 0$), then Phase 0 has a duration of zero at the start of the planning horizon ($\tau_0^{\omega} = \tau_1^{\omega} = \varphi_0^{\omega} = 0$). An example of determining the schedule for a sequence with 5 selected projects is shown in Figure 2, with given input parameters $C_{\sigma_{imp}^{i}}$, F_{ω}, ψ_{ω}, ρ,

and $\phi_{\sigma_{imp}^{i}}$. For the second and fourth project in the sequence, $\tau_{\sigma_{imp}^{i}}^{\omega} - \tau_{\sigma_{imp}^{i-1}}^{\omega} \geq \psi_{\omega}\phi_{\sigma_{imp}^{i}}$, while for other projects $\tau_{\sigma_{imp}^{i}}^{\omega} - \tau_{\sigma_{imp}^{i-1}}^{\omega} < \psi_{\omega}\phi_{\sigma_{imp}^{i}}$. After determining starting and completion times of all projects in the sequence σ_{imp}, values of these times are sorted in an ascending order, and the sorted values are assigned labels τ_1^{ω}, τ_2^{ω}, …, τ_{10}^{ω}.

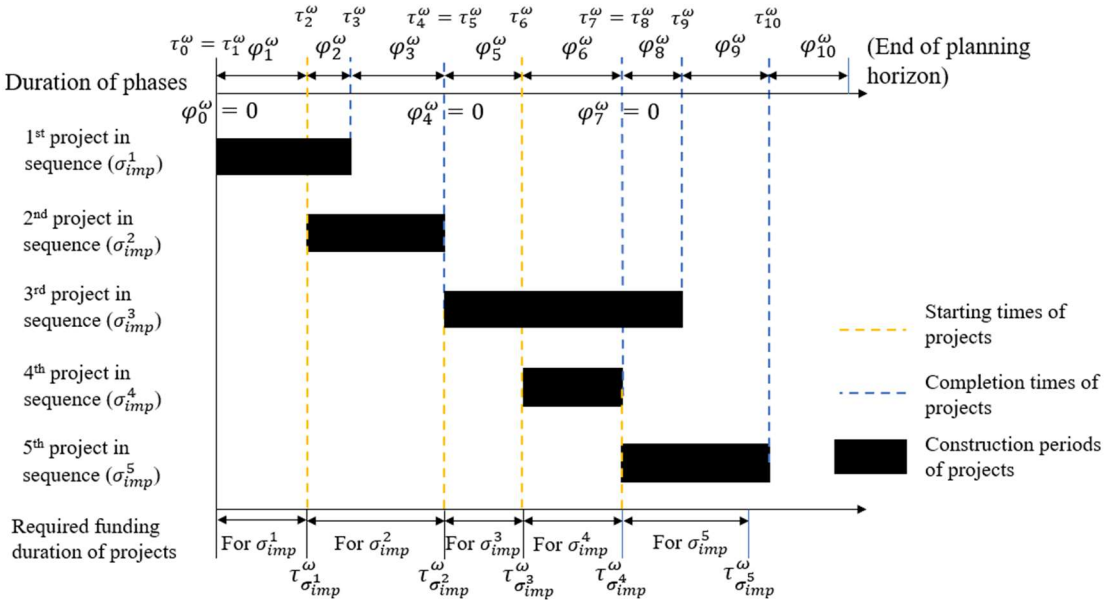


Figure 2 Determining the schedule for a sequence with 5 projects (in Problem 2)

In the example above, all projects in the sequence σ_{imp} are completed within the planning horizon under the scenario ω. In some scenarios, however, some later-started projects in the same sequence σ_{imp} cannot be completed within the planning horizon. A rule is set for truncating project starting/completion times (τ_i^{ω}) beyond the planning horizon. After determining raw values of τ_1^{ω}, τ_2^{ω}, …, $\tau_{2|\sigma_{imp}|}^{\omega}$, if a specific time value τ_j^{ω} satisfies $\tau_{j-1}^{\omega} < T \leq \tau_j^{\omega}$, then values of τ_i^{ω} with $i \geq j$ are truncated to T. This means that phase $(j-1)$ must be terminated at the end of the planning horizon, and that all phases from phase j on have zero durations. Meanwhile, values of binary

indicators $\beta^{\omega}_{\sigma^{i}_{imp}}$ in equation (19) equal to 1 by default. If a specific time value $\tau^{\omega}_{\sigma^{j}_{imp}}$ (budget-ready time for project σ^{j}_{imp}, and also the construction starting time of project σ^{j+1}_{imp}) satisfies $\tau^{\omega}_{\sigma^{j-1}_{imp}} < T \leq \tau^{\omega}_{\sigma^{j}_{imp}}$, then values of $\tau^{\omega}_{\sigma^{i}_{imp}}$ with $i \geq j$ are truncated to T, and binary indicators $\beta^{\omega}_{\sigma^{i}_{imp}}$ with $i \geq j$ equal to 0, which means that the j^{th} and later projects in the sequence cannot be paid within the planning horizon.

3.2.5 Considering choice of two modes (car and bus) in the model

The original lower-level model only considers one mode of road traffic (car) in the network. When a second mode (bus) is introduced to certain routes in the network, its impacts on equilibrium traffic flows on occupied links as well as travelers' choices between two competing modes should be considered.

In this variant of model formulation, $q_{w,P(OP)}$ refers to the hourly number of travelers (instead of vehicles) of the OD pair w during peak (off-peak) hours, and v is the unit value of travelers' time in dollars per person hour. A number ($|R|$) of bus routes R_1, R_2, …, $R_{|R|}$ are pre-determined for buses to carry passengers who travel along these routes. Each route is a simple path in the network, and is also a set of links contained in the path. The buses serving each route must perform round trips between two terminal nodes along the route, neglecting dwell times at nodes. Only the travelers whose OD pair has both its origin and destination nodes contained in the same bus route may ride a bus on this specific route. For the bus route R_i, the set of such OD pairs is denoted by W_i. It is assumed that travelers are not allowed to transfer from one bus route to another.

In all travelers of the OD pair $w \in W$, the shares of bus passengers and car users during peak (off-peak) hours are denoted as $\zeta^{bus}_{w,P(OP)}$ and $\zeta^{car}_{w,P(OP)}$, respectively. With only two modes considered:

$$\zeta^{bus}_{w,P(OP)} + \zeta^{car}_{w,P(OP)} = 1, \forall w \in W \tag{22a}$$

$$\zeta^{bus}_{w,P(OP)} = 0, \ \forall w \notin W_1, W_2, \ldots, W_{|R|} \tag{22b}$$

While car users of the same OD pair may use different paths in the network, bus passengers can only use one path along the pre-determined bus route. The passenger capacity of a bus is denoted as K_{bus}, and the average number of people in a car is denoted as K_{car}. In the peak-hour (off-peak) traffic, for each bus route R_i, the required hourly service frequency is found based on the link with the highest hourly passenger volume:

$$f_{bus,R_i} = \max_{a \in R_i} \left(\sum_{w \in W_i} q_{w,P(OP)} \zeta^{bus}_{w,P(OP)} \gamma^{w}_{a,p(R_i)} \right) / K_{bus} \tag{23}$$

where $\gamma^{w}_{a,p(R_i)}$ is a binary indicator as specified in equation (13), and $p(R_i)$ refers to the path that is completely on the bus route R_i. The resulting bus headway in hours on this route equals the reciprocal of f_{bus,R_i}:

$$h_{bus,R_i} = {f_{bus,R_i}}^{-1} \tag{24}$$

The average wait time of a passenger using route R_i can be approximated as half of the bus headway:

$$t_{wait,R_i} \approx h_{bus,R_i}/2 \tag{25}$$

Each OD pair's total hourly number of cars during peak (off-peak) hours is:

$$f^{w}_{car} = q_{w,P(OP)} \zeta^{bus}_{w,P(OP)} / K_{car} \tag{26}$$

In the computation of travel time, two possible settings of bus operation are considered: all using the dedicated bus lanes, or all sharing lanes with cars. A factor b is used for converting the number of buses to that of equivalent cars.

If buses operate on dedicated lanes, the capacity of the dedicated bus lane is assumed to be uniform throughout each route and is denoted as k_{bus,R_i} for route R_i. Based on the BPR function in equation (9), the travel time of a bus through link a on route R_i is given by:

$$t_a^{bus} = t_a\left(bf_{bus,R_i}\right) = t_a^0[1 + 0.15\frac{bf_{bus,R_i}}{k_{bus,R_i}}]^4, \forall a \in R_i \qquad (9')$$

Neglecting access and egress time, the total travel time experienced by a bus passenger of OD pair $w \in W_i$ using the bus route R_i equals in-vehicle time plus wait time. Thus, the total hourly cost of travel time of bus passengers in the peak (off-peak) network is given by:

$$Z_{P(OP)}^{bus} = \sum_{i=1}^{|R|} \sum_{w \in W_i} q_{w,P(OP)} \zeta_{w,P(OP)}^{bus} (v \sum_{a \in p(R_i) \in P_w} t_a^{bus} + v_{wait} t_{wait,R_i}) \qquad (27)$$

where v_{wait} is the unit value of passengers' wait time in \$/person hour. For the cars in the network, the BPR function and the UE traffic assignment as formulated in equations (9) to (13) are applied. The k_a in equation (9) refers to the car-dedicated capacity of each link, and the right-hand side of equation (11) is replaced by f_{car}^w. Based on equation (14), with the UE car traffic, the total hourly cost of travel time for car users in the peak (off-peak) network is given by:

$$Z_{P(OP)}^{car} = vK_{car} \sum_{a \in A} x_{a,P(OP)}^{UE} t_a\left(x_{a,P(OP)}^{UE}\right) \qquad (14')$$

Combining two modes, the total hourly cost of travel time during peak (off-peak) hours is:

$$Z_{P(OP)} = Z_{P(OP)}^{bus} + Z_{P(OP)}^{car} \tag{28}$$

With a fraction α of time being peak hours, the average total hourly cost of travel time is:

$$Z = \alpha Z_P + (1-\alpha) Z_{OP} \tag{15'}$$

If buses operate in mixed traffic with cars, then k_a in equation (9) refers to the mixed traffic capacity of each link. For link a on the bus route R_i, prior to the assignment of cars, part of its capacity k_a is occupied with the hourly bus flow of f_{bus,R_i} as found by equation (23). Keeping all the bus flows unchanged, the UE traffic assignment is performed on the cars in the network, with the right-hand side of equation (11) replaced by f_{car}^{w}. When computing the travel time through link a on the bus route R_i, the term bf_{bus,R_i} is added on top of the car flow. When the UE is reached for cars, denote the mixed traffic flow on link a in the peak (off-peak) network as $x_{a,P(OP)}^{\prime UE}$, and there are:

$$x_{a,P(OP)}^{\prime UE} = \begin{cases} x_{a,P(OP)}^{UE} + bf_{bus,R_i}, & a \in R_i, i \in 1,2,\dots,|R| \\ x_{a,P(OP)}^{UE}, & otherwise \end{cases} \tag{29}$$

In the mixed traffic, the travel time of a bus through link a on route R_i is:

$$t_a^{bus} = t_a\left(x_{a,P(OP)}^{\prime UE}\right) = t_a^0[1 + 0.15\frac{x_{a,P(OP)}^{\prime UE}}{k_{bus,R_i}}]^4, \forall a \in R_i \tag{9''}$$

The total hourly cost of travel time of bus passengers ($Z_{P(OP)}^{bus}$) is still given by equation (27), while for car users this is given by:

$$Z_{P(OP)}^{car} = vK_{car}\sum_{a\in A} x_{a,P(OP)}^{UE} t_a\left(x_{a,P(OP)}^{\prime UE}\right) \tag{14''}$$

Using equations (28) and (15'), the average total hourly cost of travel time is obtained.

In the above formulation, the mode shares $\zeta_{w,P(OP)}^{bus}$ and $\zeta_{w,P(OP)}^{car}$ are treated as parameters. However, when travelers' mode choice is considered, iteration methods provided in Section 4.4 are applied to the mode shares.

Chapter 4: Methods

4.1 Approximating effects of demand growth and discount rate

4.1.1 For the problem in rail freight networks

For the computation of monthly expected excess ($E_{\kappa_i}(\sigma_{imp})$) with long-term growth of freight demand, a simple approximation method is applied. First, the planning horizon is divided into multiple sub-periods, each of which has a duration much shorter than that of the planning horizon, so that the exponential growth rate of demand in each sub-period is approximated as being constant. If completions of improvement projects occur within a sub-period, this sub-period is further divided by these completion times. Then, using the demand level (q_w) at the midpoint of each sub-period and the corresponding network configuration, the approximated average minimized excess $X_{\omega,\kappa_i}(\sigma_{imp})$ is obtained under the disruption scenario ω from the middle-level model, and therefore the average monthly expected excess $E_{\kappa_i}(\sigma_{imp})$ in each sub-period is also obtained. In an example shown in Figure 3, the demand level grows exponentially at a constant rate. The planning horizon is first divided into five equal-length sub-periods, and then two sub-periods are further divided by the completion times of projects. Monthly expected excesses (E_{κ_i}) are computed based on approximated average demand level (q_{κ_i}) at the midpoint of each sub-period and the updated network configuration. Between two completion times of improvement projects, E_{κ_i} increases nearly proportionally with q_{κ_i}, while completion of a project tends to decrease E_{κ_i}.

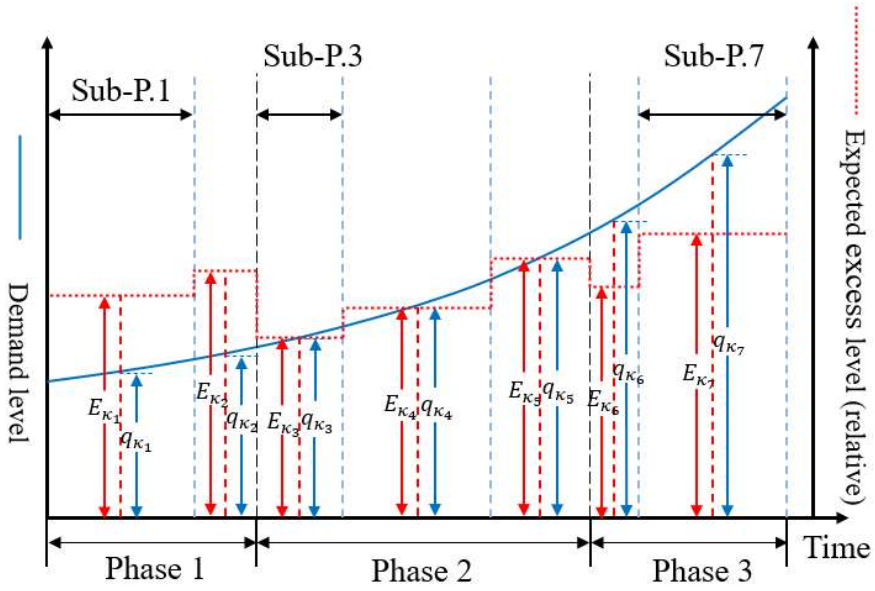


Figure 3 Approximation method with long-term demand growth using sub-periods

4.1.2 For the problem in road networks

When computing the PV of total travel time cost accumulated over the planning horizon under the improvement sequence σ_{imp} and the scenario ω, as given by $Y_{\omega,\sigma_{imp}}$ in equation (18), a similar approximation method is applied. The planning horizon is first divided into multiple equal-length sub-periods. Next, the starting and completion times of projects (τ_i^{ω}) that have been found under scenario ω and sequence σ_{imp} further divide existing sub-periods into shorter ones. After these divisions, the time τ at the midpoint of each sub-period is obtained. From the earliest to the latest, these sub-periods are denoted by ε_1, ε_2, and so forth, and their corresponding midpoint times are denoted by $\tau_{\varepsilon_1}^{\omega}$, $\tau_{\varepsilon_2}^{\omega}$, and so forth. With the scenario-specific demand growth rate g_{ω}, both peak and off-peak hourly demands ($q_{w,P(OP)}$) of all OD pairs at each midpoint time are obtained using equation (17). For each sub-period ε_i, using the midpoint

demands and the corresponding network configuration κ_{ε_i}, the approximated average hourly total travel time cost is computed using the lower-level model:

$$\bar{Z}_{UE}^{\omega,\kappa_{\varepsilon_i},\sigma_{imp}} = Z_{UE}^{\omega,\kappa_{\varepsilon_i},\sigma_{imp}}(\tau_{\varepsilon_i}^{\omega}) \tag{30}$$

The approximated cumulative travel time cost over sub-period ε_i can then be obtained by multiplying $\bar{Z}_{UE}$ with the duration of this sub-period ($\varphi_{\varepsilon_i}^{\omega}$, in years) and H (equivalent hours in a year). To obtain the PV of this cumulative travel time cost, the raw cumulative cost is divided by the discounting factor $(1+r)^{\tau_{\varepsilon_i}^{\omega}}$. Finally, $Y_{\omega,\sigma_{imp}}$ is computed by summing up PVs of cumulative travel time cost over all sub-periods ($|\varepsilon|$ is the number of sub-periods with non-zero duration in the planning horizon):

$$Y_{\omega,\sigma_{imp}} = H\sum_{i=1}^{|\varepsilon|}\frac{\varphi_{\varepsilon_i}^{\omega}\bar{Z}_{UE}^{\omega,\kappa_{\varepsilon_i},\sigma_{imp}}}{(1+r)^{\tau_{\varepsilon_i}^{\omega}}} \tag{31}$$

The iteration method for numerically finding the approximated time $\tau_{\sigma_{imp}^{i}}^{\omega}$ (the budget-ready time for the i^{th} project in sequence σ_{imp} under scenario ω) is presented as follows. Given a known value of $\tau_{\sigma_{imp}^{i-1}}^{\omega}$, the first tentative value of $\tau_{\sigma_{imp}^{i}}^{\omega}$ equals the one given by equation (20a) where only external budget supply is available. This tentative value, denoted by $\hat{\tau}_{\sigma_{imp}^{i}}^{\omega}$, is larger than the one given by (20b) with internal budget supply available. Then, all the resulting sub-periods between $\tau_{\sigma_{imp}^{i-1}}^{\omega}$ and the tentative $\hat{\tau}_{\sigma_{imp}^{i}}^{\omega}$ are found. Using the hourly demands at the midpoint of each sub-period and its corresponding network configuration, the approximated average hourly total travel time cost ($\bar{Z}_{UE}^{\omega,\kappa_{\varepsilon_i},\sigma_{imp}}$) is obtained. With this approximation, in each sub-period ε_i in between $\tau_{\sigma_{imp}^{i-1}}^{\omega}$ and $\hat{\tau}_{\sigma_{imp}^{i}}^{\omega}$, available budget is accumulated at a constant rate of

$(F_\omega + \rho H \bar{Z}_{UE}^{\omega,\kappa_{\varepsilon_i},\sigma_{imp}})$ per year. Thus, an increment of time from $\tau_{\sigma_{imp}^{i-1}}^{\omega}$, denoted by τ_Δ, can be found so that:

$$C_{\sigma_{imp}^{i}} = \sum_{j=j_1}^{j_2} (F_\omega + \rho H \bar{Z}_{UE}^{\omega,\kappa_{\varepsilon_j},\sigma_{imp}}) \cdot \min\left\{\varphi_{\varepsilon_j}^{\omega}, \max\{\tau_{\sigma_{imp}^{i-1}}^{\omega} + \tau_\Delta - \tau_{\varepsilon_j}^{\omega} + \varphi_{\varepsilon_j}^{\omega}/2, 0\}\right\}, j \geq 1 \quad (32)$$

subject to:

$$\tau_{\varepsilon_{j_1-1}}^{\omega} < \tau_{\sigma_{imp}^{i-1}}^{\omega} < \tau_{\varepsilon_{j_1}}^{\omega}, \qquad 2 \leq i \leq |\sigma_{imp}| \quad (33a)$$

$$\tau_{\varepsilon_{j_2}}^{\omega} + \varphi_{\varepsilon_{j_2}}^{\omega}/2 = \hat{\tau}_{\sigma_{imp}^{i}}^{\omega}, \qquad 1 \leq i \leq |\sigma_{imp}| \quad (33b)$$

$$0 < \tau_\Delta < \hat{\tau}_{\sigma_{imp}^{i}}^{\omega} - \tau_{\sigma_{imp}^{i-1}}^{\omega}, \qquad 1 \leq i \leq |\sigma_{imp}| \quad (33c)$$

$$\tau_{\sigma_{imp}^{0}}^{\omega} = 0 \quad (33d)$$

The resulting value of $(\tau_{\sigma_{imp}^{i-1}}^{\omega} + \tau_\Delta)$ is the approximated time (in years) when the available funding becomes exactly sufficient for the i^{th} project in sequence σ_{imp} under scenario ω. This time value is denoted by τ_I. Then, the sub-period (time interval) is found that contains τ_I, and the ending time of this sub-period is denoted by τ_{II}. The "real" budget-ready time $\tau_{\sigma_{imp}^{i}}^{\omega}$ as estimated with midpoint demand levels is located between τ_I and τ_{II}. Note that when the budget-ready time $\tau_{\sigma_{imp}^{i}}^{\omega}$ changes, the duration of the sub-period that ends at $\tau_{\sigma_{imp}^{i}}^{\omega}$ also changes. As a result, for each trial value of $\tau_{\sigma_{imp}^{i}}^{\omega}$, the value of $\bar{Z}_{UE}^{\omega,\kappa_{\varepsilon_j},\sigma_{imp}}$ corresponding to the sub-period ending at $\tau_{\sigma_{imp}^{i}}^{\omega}$ must be updated.

The search for the "real" budget-ready time $\tau_{\sigma_{imp}^{i}}^{\omega}$ uses a bisection method. Let τ_M be the average of τ_I and τ_{II}. τ_M serves as the temporary value of $\tau_{\sigma_{imp}^{i}}^{\omega}$. Among all sub-

periods between $\tau_{\sigma_{imp}^{i-1}}^{\omega}$ and τ_M, the latest one ends at the time τ_M. With these updated sub-periods (including updated values of $\varphi_{\varepsilon_j}^{\omega}$ and $\bar{Z}_{UE}^{\omega,\kappa_{\varepsilon_j},\sigma_{imp}}$), the following item Δ is computed:

$$\Delta = C_{\sigma_{imp}^{i}} - \sum_{j=j_1}^{j_2} (F_\omega + \rho H \bar{Z}_{UE}^{\omega,\kappa_{\varepsilon_j},\sigma_{imp}}) \cdot \varphi_{\varepsilon_j}^{\omega}, \qquad j \geq 1 \tag{34}$$

subject to:

$$\tau_{\varepsilon_{j_1-1}}^{\omega} < \tau_{\sigma_{imp}^{i-1}}^{\omega} < \tau_{\varepsilon_{j_1}}^{\omega}, \qquad 2 \leq i \leq |\sigma_{imp}| \tag{35a}$$

$$\tau_{\varepsilon_{j_2}}^{\omega} + \varphi_{\varepsilon_{j_2}}^{\omega}/2 = \tau_M < \hat{\tau}_{\sigma_{imp}^{i}}^{\omega}, \qquad 1 \leq i \leq |\sigma_{imp}| \tag{35b}$$

$$\tau_{\sigma_{imp}^{0}}^{\omega} = 0 \tag{35c}$$

If $\Delta > 0$, let τ_M be the new τ_I in the next iteration. If $\Delta < 0$, let τ_M be the new τ_{II} in the next iteration. An updated $\tau_M = (\tau_I + \tau_{II})/2$ becomes the new temporary value of $\tau_{\sigma_{imp}^{i}}^{\omega}$, which starts the next iteration. The iteration process loops until the absolute value of Δ becomes smaller than a predetermined threshold. Upon termination, τ_M becomes the final value of $\tau_{\sigma_{imp}^{i}}^{\omega}$.

In the search of $\tau_{\sigma_{imp}^{i}}^{\omega}$ based on existing $\tau_{\sigma_{imp}^{i-1}}^{\omega}$, if for a specific j both $\hat{\tau}_{\sigma_{imp}^{j}}^{\omega}$ and the first value of τ_I exceed T (the duration of the planning horizon), then the numerical search is terminated. For all $i \geq j$, values of $\tau_{\sigma_{imp}^{i}}^{\omega}$ are set to T, and the binary indicators $\beta_{\sigma_{imp}^{i}}^{\omega}$ equal 0, which means that the j^{th} and later projects in the sequence cannot be paid within the planning horizon.

4.2 Traffic assignment with congestion

For the problem in road networks, the congestion function of each link is given by equation (9). To attain user equilibrium as specified in the nonlinear programming problem in (10) to (13) and then compute Z_{UE}, the Frank-Wolfe (F-W) algorithm, proposed by Frank and Wolfe (1956), is used and customized for this problem. The steps of this algorithm are shown below.

0) Initialization. Let flows on all links be zero, and do all-or-nothing traffic assignment based on the free-flow travel times of links (that is, for each OD pair w, find out the path p with the shortest travel time, and let $f_p^w = q_w$ for this path). Compute the resulting traffic flow on each link and obtain the flow vector $\{x_a^1\}$. This is followed by the start of iteration i=1.

1) At the start of iteration i, compute the travel time vector $\{t_a^i\}$ where $t_a^i = t_a(x_a^i)$ for each link. Then do all-or-nothing assignment based on $\{t_a^i\}$, and compute the resulting traffic flow to obtain the auxiliary flow vector $\{u_a^i\}$, which serves as an updated searching direction of traffic flows.

2) Find the optimal step length $\lambda_i \in [0,1]$ that will move the flow vector from $\{x_a^i\}$ to $\{x_a^{i+1}\}$. A bisection method is used to search for λ_i that satisfies the following equation:

$$\sum_{a \in A} (u_a^i - x_a^i) \cdot t_a[x_a^i + \lambda_i(u_a^i - x_a^i)] = 0 \tag{36}$$

3) Obtain the succeeding flow vector $\{x_a^{i+1}\}$ where $x_a^{i+1} = x_a^i + \lambda_i(u_a^i - x_a^i)$. Compute the relative gap between flow vectors $\{x_a^i\}$ and $\{x_a^{i+1}\}$ given by:

$$\Delta_x = \sqrt{\sum_{a \in A} (x_a^{i+1} - x_a^i)^2} \Big/ \sum_{a \in A} x_a^i \tag{37}$$

4) If the relative gap is smaller than a pre-specified threshold, stop the iterative search and use $\{x_a^{i+1}\}$ as the final vector of UE traffic flow. Otherwise, loop back to step 1) and perform the next iteration.

It should be noted that F-W algorithm provides a relatively rough approximation of UE traffic flows. For a more precise approximation, a mesoscopic simulation method such as cell transmission or link transmission may be used for assisting the UE traffic assignment.

4.3 Genetic algorithm for project selection and sequence optimization

In the short-term restoration in Problem 1, with n damaged components in the network, there are $n!$ possible sequences for restoration. In the long-term improvement in both problems, with n candidate projects, the number of all possible selection and sequences of projects exceeds $n!$. If an exhaustive enumeration is used for finding the exact optimal solution, the computation time will be at least $O(n!)$. As n increases, this computation time quickly becomes unacceptable for practical applications. Therefore, a heuristic method is proposed for delivering an approximately optimal solution in a reasonable time, and a genetic algorithm (GA) is customized and applied in this dissertation.

The following descriptions of the GA are based on the short-term restoration optimization problem in a rail freight network. When optimizing the long-term selection and scheduling network improvement projects in both types of networks, this

GA can be applied in a similar manner to short-term restoration optimization, with some differences specified in section 4.3.3.

4.3.1 Representation and evaluation of sequences

In this GA, each possible restoration sequence under a given disruption scenario is encoded into a "chromosome". In a disruption scenario, m nodes and n links are damaged. The m damaged nodes are first sorted by numeric labels and then assigned new labels (for damaged components) from 0 to (m-1). The n damaged links are sorted by alphabetic labels and assigned labels from m to (n-1). There are (m+n)! possible permutations using these new labels, and each permutation written inside parentheses forms a "chromosome", which represents a unique restoration sequence. For example, if nodes labeled 6, 2, 4 and links labeled d, g, b are damaged, then the components labeled 2, 4, 6, b, d, g can be assigned with labels 0, 1, 2, 3, 4, 5, respectively. A chromosome written as (2, 3, 5, 1, 4, 0) represents a restoration plan that sequentially restores node 6, link b, link g, node 4, link d, and node 2.

Evaluation of a chromosome is straightforward. Each chromosome's fitness value is simply the excess incurred during the corresponding restoration process, which can be computed using the method specified above. A smaller fitness value means a better chromosome, i.e., a lower cost plan. It should be noted that, a chromosome with (m+n) locations (numbers) involves (m+n) different intermediate network states, each (except the first) having a capacity update by one component with respect to its previous state. Every time capacity settings change in a network, shipments may need to be reassigned in order to minimize the total hourly cost. When multiple chromosomes are evaluated, while they share a common network state at the first restoration step, each of their later

steps can experience different intermediate network states. The kth step of a restoration plan has C_{m+n}^{k-1} possible intermediate states. To avoid repetitive evaluation of a network state (using the LP) and reduce computation time, at each step of each restoration sequence the algorithm checks whether an intermediate network state has been previously evaluated by looking for this state in a list. If yes, this state's minimized hourly cost is directly extracted without running the LP. Otherwise, the LP is run, and this state is recorded in the list along with its corresponding minimized hourly cost.

4.3.2 GA operators

4.3.2.1 Population initialization

In the problem of short-term restoration sequencing in rail freight networks, this GA starts with an initial population of *pop_size* (an even number between 20 and 40) chromosomes. A large fraction (over 80%) of these chromosomes are generated by randomly arranging integers 0 to (m+n-1), which is equivalent to drawing (m+n) times from integers 0 to (m+n-1) without replacement. For the rest (less than 20%), a strategy is used for providing potentially better starting points for chromosome searching: looking for the "recommended first step" of restoration. Each of (m+n) possible first steps has its own "first step restoration time", during which the network must be in the disrupted state without restoration and a minimized hourly cost increment denoted as the "initial hourly increment" is incurred. After completion of each possible first step, the network enters its corresponding intermediate state, where only one component is restored and a minimized hourly cost increment denoted as "next-step hourly increment" is incurred. Noting that all possible restoration sequences have the same "total restoration time" (i.e., the sum of all restoration time excluding access time), an

upper bound of total cumulative cost increment (excess) can be computed for each possible first step by: “first step restoration time” × (“initial hourly increment” – “next-step hourly increment”) + “total restoration time” × “next-step hourly increment”. The “recommended first step” of restoration yields the smallest upper bound. When generating less than 20% of the initial population, the first number in a chromosome represents the “recommended first step”, while the remaining (m+n-1) numbers are randomly arranged.

4.3.2.2 Parent selection

For each generation thereafter, GA operators (selection, crossover, mutation) are applied to evaluated chromosomes so that its next generation is generated. This iteration continues until the best fitness value in a generation remains unimproved for a certain number (denoted as *max_stall*) of generations or the maximal iteration count (denoted as *max_iter*) is reached.

After evaluating all chromosomes in a generation, “parent” chromosomes are selected for generating “children” in the next generation. Before selection, the fitness values in this generation are sorted in ascending order, and directly succeed a certain number (denoted as *best_chroms,* usually an even number smaller than 20% of *pop_size*) of chromosomes with lowest (best) fitness values to the next generation. Then, the number of “children” needed in the next generation equals *pop_size* minus *best_chroms*. The same number of “parents” are selected in the current generation. Each chromosome may be selected multiple times and some “parents” may be duplicated. Chromosomes with lower fitness values (i.e., lower costs) are given higher probabilities of being selected as “parents”. The selection probabilities are based on the fitness rankings of chromosomes, so that the selective pressure stays constant over

generations and is not affected by absolute differences of fitness values (Whitley, 1989). Let the chromosome with the lowest fitness value have the ranking value of 1. Also let i be the ranking value (an integer between 1 and *pop_size*) of a chromosome in the current generation. The selection probability of this chromosome is given by:

$$(Prob)_i = \frac{sel\_pres(1 - sel\_pres)^{i-1}}{1 - (1 - sel\_pres)^{pop\_size}} \tag{38}$$

where $0 < sel\_pres < 1$. A greater sel_pres poses greater selective pressure.

In each selection operation, two different "parent" chromosomes are selected from the current generation with their corresponding probabilities. After potential crossover and mutation, they produce two "children" for the next generation. These two "parents" are replaced into the population for the next selection. The operation loop of selection-crossover-mutation is executed (*pop_size* - *best_chroms*)/2 times until the number of chromosomes in the next generation reaches *pop_size*.

4.3.2.3 Crossover and mutation operators

The crossover operator deals with two "parents" at a time and produces two "children". The probability that crossover between two "parents" actually occurs is given by a parameter p_c. Before each crossover operation, a number uniformly distributed between 0 and 1 is randomly generated. Crossover actually occurs only if this number is smaller than p_c. Otherwise, no crossover occurs and "children" are identical to their "parents" before possible mutation. In a crossover operation, segments of two chromosomes are swapped. In a chromosome with (m+n) integers, two different locations are randomly chosen from (m+n+1) possible locations (those between numbers plus two ends). For each of two chromosomes to be operated, the segment between these two locations is swapped with the corresponding segment in the other

chromosome. This may create infeasible chromosomes with duplicate integers (labels for disrupted components), and the Partial Mapped Crossover (PMX) method, proposed by Goldberg and Lingle (1985), is applied to fix the error.

"Child" chromosomes may experience mutation before they are finally transmitted to the next generation. The probability that mutation of a "child" actually occurs is given by a parameter *p_m*. If a number randomly generated from U(0,1) is smaller than *p_m*, mutation actually occurs. Given that a mutation occurs, two numbers (whose locations in a chromosome are denoted as *loc1* and *loc2*, with *loc1*<*loc2*) are randomly chosen from (*m*+*n*) numbers in a chromosome. With a probability of 0.5, the operator simply swaps these two numbers. With the remaining probability of 0.5, the operator performs an insertion, which moves one number from its previous location in the chromosome to a new one. With equal probability, it either removes the number at *loc1* and inserts it just before *loc2*, or takes out that at *loc1* and inserts it just after *loc2*.

4.3.3 The GA in upper-level problems

The optimization of the long-term improvement plan in the upper-level model (in both Problem 1 and Problem 2) uses a similar procedure. For problems with 6 or fewer improvement projects, exhaustive enumeration is used. Otherwise, a similar GA is applied, with some modifications in chromosome representation and operator settings. In population initialization, the "recommend first step" strategy is not applied. The length of a chromosome equals the number of candidate projects ($|\Lambda|$). After generating each chromosome by randomly arranging integers 0 to ($|\Lambda|$-1) (or 1 to $|\Lambda|$), for each location in the chromosome there is a specified probability that the number is replaced with a blank. When evaluating each chromosome, blank locations are removed, and the

remaining sequence of numbers represents the implementation sequence of improvement projects. A smaller objective function value (as given by equation (6) in Problem 1 and equation (19) in Problem 2) with respect to an improvement sequence σ_{imp} means a better fitness value of its corresponding chromosome. Blank locations participate in crossover and mutation operators and behave similarly to numbers. In the crossover operator, the blank locations are temporarily filled with numbers so that each integer from 0 to (|Λ|-1) (or 1 to |Λ|) appears in each chromosome exactly once. When segments in two chromosomes are swapped and the PMX method is performed, these labeled blank locations move together with their corresponding numbers. At the end of crossover, these blank locations in the two chromosomes return blank. In the mutation operator, with a given probability, the action taken is to randomly select a location in the chromosome being processed. If the location has a number, it mutates into a blank, and vice versa (while it should be ensured that each project code appears only once in a chromosome). With the remaining probability, the operator takes the action of swapping or insertion equiprobably, as specified in Section 4.3.2. For each improvement sequence to be evaluated, the algorithm checks whether it has been previously evaluated by looking for it in a list. If yes, its corresponding objective function value is directly retrieved from the list to avoid repetitive computation. Otherwise, the objective function value is computed for this "new" sequence and recorded in the list along with the sequence.

4.4 Iteration of mode shares when bus and car are considered

In Problem 2, when two modes – bus and car – are considered, the mode shares $\zeta_{w,P(OP)}^{bus}$ and $\zeta_{w,P(OP)}^{car}$ for the OD pair w in the peak (off-peak) network influence traffic

flows and their resulting travel times. For travelers of the same OD pair that is covered by one of the bus routes (i.e., $w \in W_i$), the difference in the impedance to travel by bus or by car will influence their mode choices, causing a shift in the mode shares.

Let $\zeta^{bus}_{w,P(OP)}$ and $\zeta^{car}_{w,P(OP)}$ be input parameters to the formulation in Section 3.2.5, and the outputs are used for computing travelers' impedances by the two modes. For the OD pair $w \in W_i$, the impedance for each bus passenger equals his travel time cost plus bus fare (c_{bus}):

$$I^{bus}_w = v \sum_{a \in p(R_i) \in P_w} t^{bus}_a + v_{wait} t_{wait,R_i} + c_{bus}, \quad \forall w \in W_i \tag{39}$$

where t_{wait,R_i} is the output from equation (25), and t^{bus}_a is the output from equation (9') if buses use dedicated lanes or from equation (9'') if buses share traffic with cars. The impedance for each car user using the path $p \in P_w$ equals his travel time cost plus fuel cost:

$$I^{car}_{w,p} = v t^{car}_p + c_{fuel} d_p / K_{car}, \quad \forall p \in P_w, w \in W_i \tag{40}$$

where t^{car}_p is the travel time through path p by car, given by $\sum_{a \in p} t_a\left(x^{UE}_{a,P(OP)}\right)$ if buses use dedicated lanes (see equation (14')) or $\sum_{a \in p} t_a\left(x'^{UE}_{a,P(OP)}\right)$ if buses share traffic with cars (see equation (14'')), d_p is the length of path p, given by $\sum_{a \in p} d_a$, and c_{fuel} is the unit car fuel cost in dollar per vehicle mile. Among the car users of the OD pair w, the fraction of them using path p is denoted by χ^w_p. For the OD pair w, the weighted average of impedance by car is:

$$I^{car}_w = \sum\nolimits_{p \in P_w} \chi^w_p \, I^{car}_{w,p}, \quad \forall w \in W_i \tag{41}$$

To determine the values of χ_p^w, the "step lengths" during the iterations of the F-W algorithm (see Section 4.2) are considered. For example, when the algorithm is started, all car traffic of the OD pair w is assigned to path p_1, making $\chi_{p_1}^w = 1$. In iteration 1 the optimal step length found is $\lambda_1 = 0.3$, with an all-or-nothing assignment to path p_2. After iteration 1, the fractions become $\chi_{p_1}^w = 1 - 0.3 = 0.7$, and $\chi_{p_2}^w = 0.3$. In a more general case, before iteration i, the fractions are $\chi_{p_1}^w$, $\chi_{p_2}^w$, $\chi_{p_3}^w$, …, the sum of which being 1. After iteration i, with the optimal step length λ_i and an all-or-nothing assignment to path p_j, the fractions become $\chi_{p_j}^{\prime w} = \chi_{p_j}^w + \lambda_i(1 - \chi_{p_j}^w)$, and $\chi_{p_k}^{\prime w} = \chi_{p_k}^w(1 - \lambda_i)$ for all $k \neq j$. When the algorithm terminates, the values of χ_p^w are determined.

With the logit mode choice model, the resulting mode shares are given by:

$$\zeta_{w,P(OP)}^{\prime bus} = \frac{e^{-I_w^{bus}}}{e^{-I_w^{bus}} + e^{-I_w^{car}}}, \qquad \zeta_{w,P(OP)}^{\prime car} = 1 - \zeta_{w,P(OP)}^{\prime bus}, \qquad \forall w \in W_i \qquad (42)$$

These output mode shares become new inputs in the next iteration. As iterations continue, the share of each mode is expected to converge. The iteration of mode shares terminates when the maximum iteration count is reached or the largest absolute difference between input and output shares among all OD pairs does not exceed a preset threshold. The final output mode shares are used as the finalized inputs for computing hourly total travel time cost at the lower-level model.

Chapter 5: Numerical results for Problem 1

5.1 A small example network and its parameters

The following numerical case, along with the LP formulation for hourly cost minimization and the simple GA for optimizing short-term restoration, are coded in Python 3.7.3. The program is run on a personal laptop with an Intel® Core™ i7-8750H CPU @ 2.20GHz. The DOcplex library is used for formulating and solving LP problems, while the NetworkX library is used for generating the small test network with attributes (capacity and length) attached to nodes and arcs. Shortest paths, simple paths, and their corresponding travel times by rail are obtained using NetworkX.

To demonstrate this integrated optimization model for short-term and long-term resilience of rail freight network, a small network with 10 nodes and 13 links is synthesized. These components are connected as shown in Figure 4, with each two-directional link labeled with its length and capacity in each direction. Each link corresponds to two directed arcs in the arc set A. The capacities of nodes and hourly amounts of freight that must be delivered for each OD pair are listed in Table 4.

The model is first demonstrated at the lower and middle level. For various cost components, the unit costs are c_r=\$0.5/ton/mi for rail shipment, c_u=\$10/tons/hr for users' time value, and c_{alt}=5c_r=\$2.5/ton/mi for shipment by an alternative mode (trucks are assumed here). Both normal train speed and truck speed are 40 mi/hr. For each OD pair $w \in W$ with non-zero demand, the truck route is assumed to follow the shortest rail path in the network, whose length and normal travel time are denoted as d_w and t_w, respectively. The OD-pair-specific unit truck shipment cost c_{alt}^w in \$/ton is

given by $c_{alt}^{w} = c_{alt}d_w + c_u t_w$, and the maximum allowable rail shipment time T_w is assumed to be $2.5t_w$.

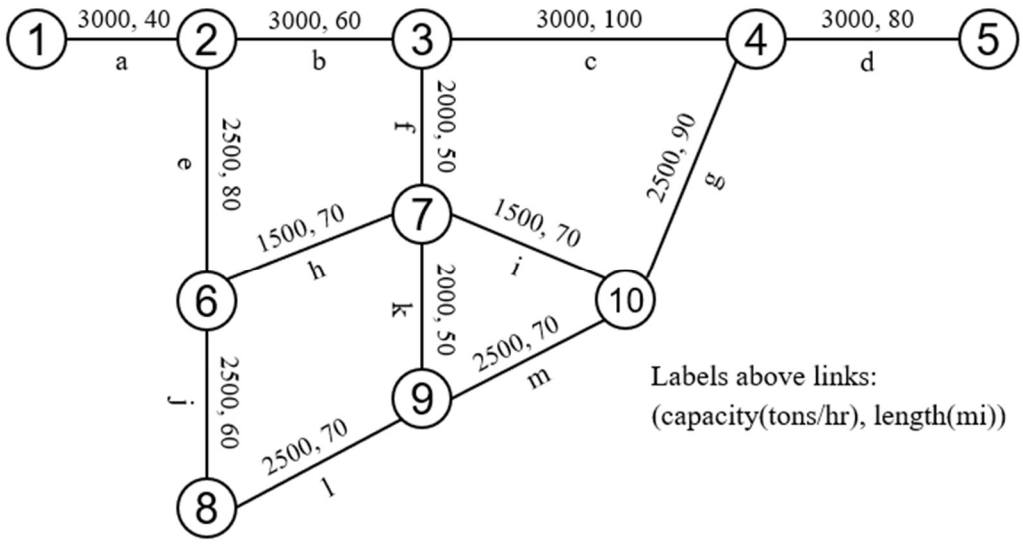


Figure 4 The small example rail freight network

Table 4 Capacities of nodes and the demand matrix

Node Label	Capac. (t/hr)		O-D Dem. (t/hr)	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
1	6,000		O1					800				700	
2	6,000		O2				600						1,200
3	6,000		O3		600							800	
4	6,000		O4							900	700		
5	6,000		O5		800				900				
6	5,000		O6	500									900
7	7,000		O7					600			1,000		
8	5,000		O8	600						800			
9	5,000		O9			500			700				
10	6,000		O10			900	800						

5.2 Optimized short-term restoration sequences and schedules

5.2.1 Results for an example disruption scenario

In an example scenario, components with dashed lines are damaged, as shown in Figure 5. It takes 10 hours to complete restoration of each damaged node, while each link is restored at a rate of 10 miles per hour. In the disrupted network state without

restored components, capacities of damaged components are halved, and travel times through damaged links are doubled. All work teams depart from the base node 7 and travel at 40 miles per hour.

GA parameters are set as follows: *pop_size* = 30, *best_chroms* = 2, *max_iter* = 150, *max_stall* = 30, *p_c* = 0.5, *p_m* = 0.8, *sel_pres* = 0.06. In the initial population 3 chromosomes are generated with the "recommended first step". Components 3, 9, b, c, f, k, l, m are re-labeled with integers 0 to 7, respectively, for generating chromosomes. Given this disruption scenario, if there is only one work team available for restoration, then the GA-optimized restoration sequence is represented as the chromosome (0, 2, 4, 1, 6, 5, 3, 7). This is the optimized result in all 10 runs with different initial populations and an average runtime of 4.816 seconds. The corresponding restoration sequence that minimizes the excess during the restoration process is:

Node 3 → Link b → Link f → Node 9 → Link l → Link k → Link c → Link m

With this optimized sequence, the work team's route in this network is marked in Figure 5.

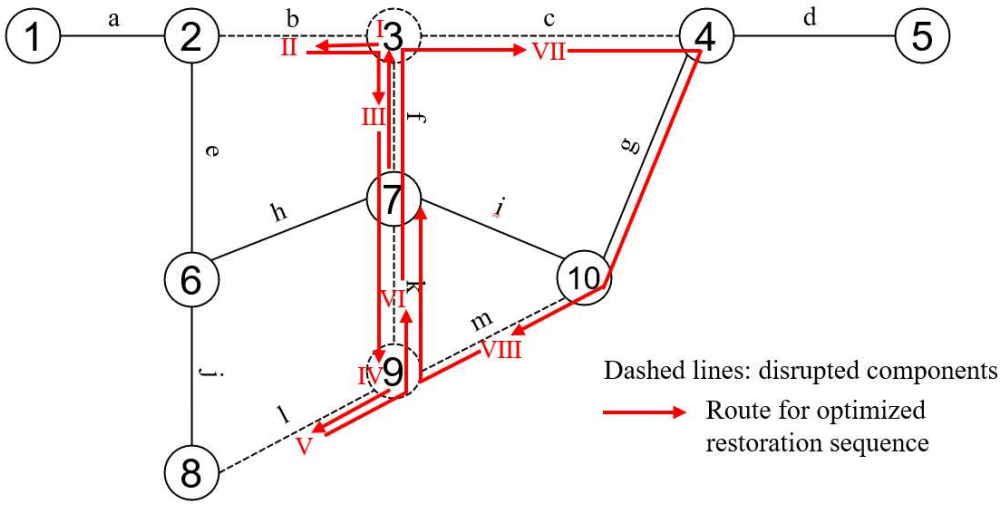


Figure 5 Optimized restoration route under the example disruption scenario

Figure 6 shows the restoration schedule of the work team, which is uniquely determined by the optimized restoration sequence, given the access times as well as the restoration durations for damaged components. Changes in minimized hourly cost during the restoration process are also shown in Figure 6. The minimized hourly cost in normal network operation is $2,205,500, and the cumulative cost increment (excess) beyond normal operation cost, as shown by the shaded area, is $28,619,062.5. An exhaustive enumeration performed on all 8!=40,320 possible restoration sequences reveals that this optimized sequence does have the globally lowest excess. This minimized excess can be used as a measure of short-term resilience for this example network and disruption scenario with one work team.

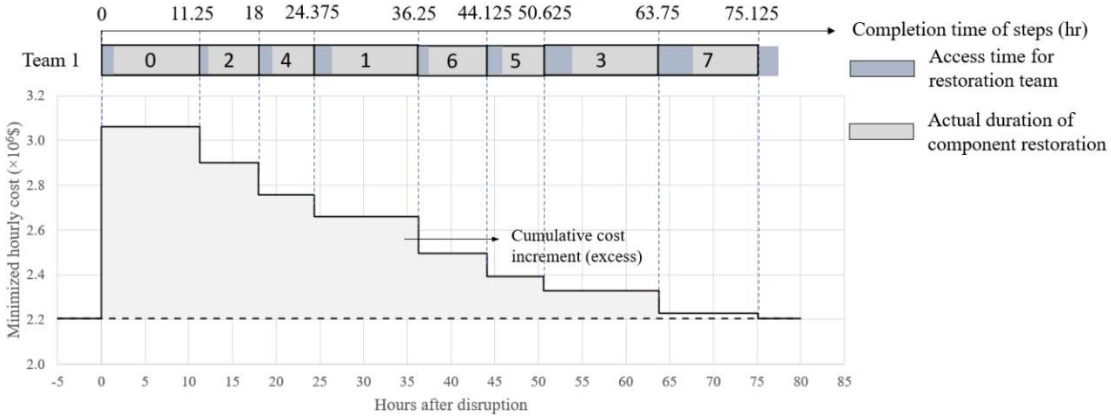


Figure 6 Optimized restoration schedule and corresponding hourly cost changes

With other parameters unchanged, the excess is also minimized when there are 2 and 3 work teams available. The optimized restoration sequences with 2 and 3 teams are represented as chromosomes (0,1,2,5,6,4,3,7) and (1,0,4,2,6,3,5,7), respectively. The uniquely determined restoration schedules and their resulting changes in minimized hourly cost are shown in Figure 7. Minimized excesses are $16,985,710.9 and

$12,759,437.5 with 2 and 3 teams, respectively, with their global optimality verified through exhaustive enumeration.

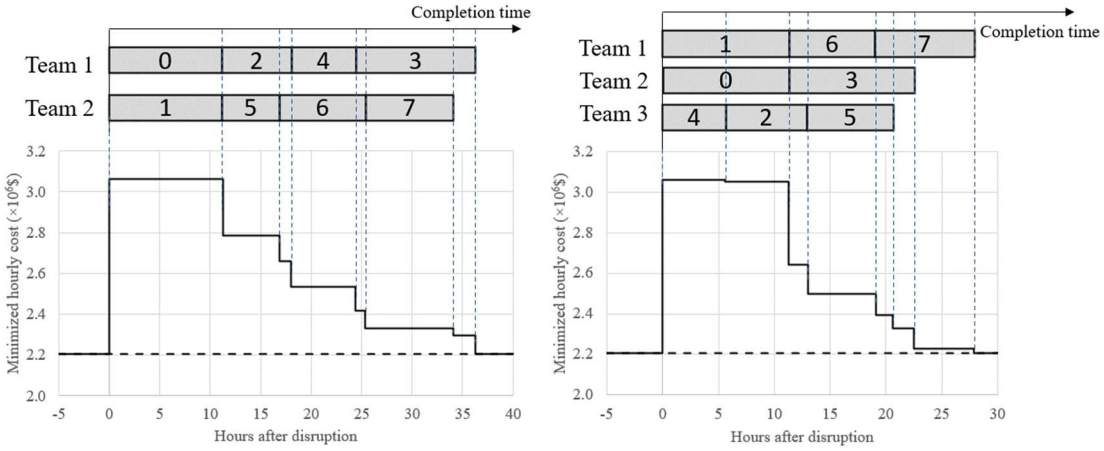


Figure 7 Optimized schedule & corresponding hourly cost changes with more teams

5.2.2 Testing GA with more disruption scenarios

Other disruption scenarios are set up to test the GA. For each scenario the model is run 10 times. Results are summarized in Table 5. Numbers following a "*" denote numbers of each distinct observation in 10 runs. For the nth chromosome listed in the fourth row, its corresponding excess and global optimality are respectively the nth item listed in following rows as well. Global optimality of the GA solutions is verified here through exhaustive enumeration for each scenario. It can be learned that, within acceptable computation time, the GA usually finds the globally optimal restoration sequences and sometimes the second optimal ones, whose excesses are less than 0.1% above the lowest ones. These excesses of GA solutions can be treated as measures of short-term network resilience under these additional tested disruption scenarios.

Table 5 GA results under additional disruption scenarios

Damaged nodes	6, 8, 9	2, 3, 6	3, 4, 10
Damaged links	e, h, j, k, l, m	a, b, e, h, j	b, c, d, f, g, i, m

# damaged components	9	8	10
Avg. GA comp. time (s)	9.698	5.419	16.263
Exhaustive enumeration comp. time (s)	27.771	5.160	384.958
Optimized restoration sequence by GA	(6,2,0,7,1,5,3,4,8) *3 (4,0,1,5,2,7,6,3,8) *7	(1,4,0,2,7,5,3,6) *10	(1,0,4,6,8,2,5,3,7,9) *9 (1,0,6,4,5,8,2,3,7,9) *1
Cumulative cost increment (excess) ($)	26,438,375 * 3 26,462,125 * 7	30,112,062 * 10	51,733,062 * 9 51,774,312 * 1
Global solution optimality	Best * 3 Second best * 7	Best * 10	Best * 9 Second best * 1

5.3 Optimized long-term improvement of the network

5.3.1 Sets of disruption scenarios and improvement projects

For the same initial network configuration with the demand matrix unchanged, there are 7 synthesized disruption scenarios, each having a center of disruption. When a scenario is centered at a node, the node itself, links connecting this node, and outer nodes connected by these links have their capacities reduced by 70%, 40%, and 20%, respectively. If a link becomes the center, then capacities also decrease by 70%, 40%, and 20% for this link, its two end nodes, and the outer links connected to these two nodes, respectively. If the capacity of a damaged link is x (x<1) times of the original, then the travel time through this link is assumed here to be 1/x times of the original. The centers of disruption in these 7 scenarios are Node 2, Node 4, Node 7, Node 8, Link f, Link g, and Link j, with monthly occurrence probabilities (π_ω) of 0.015, 0.02, 0.015, 0.01, 0.025, 0.015, 0.015, respectively.

In a planning horizon with a specified duration, 5 projects are available for long-term network improvement. With a constant external budget supply of $\$5\times10^6$ per month, details of these projects are shown in Table 6.

Table 6 Candidate improvement projects and their parameters in the smaller rail network

No.	Action of project	Construction cost C_l(× \$$10^6$)	Budget accumulation time (months)	Minimum work duration (months)
1	Increase capacities of Links a, b, c, d and Nodes 1, 2, 3, 4, 5 by 1/3	110	22	24
2	Increase capacities of Links e, g, j, l, m and Nodes 6, 8, 9, 10 by 40%	164	32.8	30
3	Increase capacities of Links f, h, i, k and Node 7 by 40%	100	20	24
4	Add a 90-mile link with 4000 tons/hr capacity, connecting 1 and 6	135	27	24
5	Add a 120-mile link with 3000 tons/hr capacity, connecting 4 and 7	180	36	30

In the computation of minimized hourly cost, unit cost parameters (c_r, c_u, and c_{alt}) in this upper-level numerical case are doubled from those in the lower-level case. All other parameters are unchanged from those used in the lower-level case.

5.3.2 The optimized selection and schedule of improvement with integrated short-term results

With only 5 projects, the optimal long-term improvement plan can be found through exhaustive enumeration in acceptable computation time (<30min). If there is only one work team, and the planning horizon lasts 300 months, then the optimal plan is to sequentially complete projects #1 and #2 (as shown in Table 6). Completion times of projects #1 and #2 are 24 and 54.8 months into the planning horizon, respectively. Therefore, in the planning horizon phases 0, 1, and 2 last 24, 30.8, and 245.2 months, respectively. The resulting total cost (TC, cumulative expected excess plus total construction cost) during the planning horizon is \$11.2226×10^8, which is more than \$100 million below the \$12.3413×10^8 without improvement projects. With this optimal long-term improvement plan, Table 7 shows the underlying optimized short-term

restoration sequences and their corresponding minimized excesses under each disruption scenario in different phases.

Table 7 Underlying short-term results with optimal long-term improvement

Phase No.	Center of disruption in this scenario	Optimized restoration sequence	Minimized excess ($)	Monthly expected excess ($)
0	Node 2	2→b→3→6→e→a→1	40,190,625.0	4,113,762.5
	Node 4	4→3→c→d→g→10→5	52,025,996.5	
	Node 7	7→i→3→f→6→9→k→h→10	50,134,520.8	
	Node 8	8→6→9→l→j	22,076,687.5	
	Link f	7→f→3→i→c→b→k→h	29,406,984.4	
	Link g	4→g→10→i→c→d→m	33,975,640.6	
	Link j	6→j→8→l→e→h	16,852,619.8	
1	Node 2	2→b→3→a→e→6→1	36,523,791.7	3,582,761.9
	Node 4	4→3→c→d→g→10→5	50,422,041.7	
	Node 7	7→3→f→1→k→9→6→h→10	43,012,472.2	
	Node 8	8→6→j→l→9	15,883,375.0	
	Link f	3→f→7→i→b→c→k→h	25,385,580.7	
	Link g	4→g→10→i→c→d→m	28,322,203.1	
	Link j	6→j→8→l→e→h	10,864,718.8	
2	Node 2	2→b→3→e→a→6→1	30,971,125.0	2,606,557.4
	Node 4	4→c→3→d→g→5→10	45,558,416.7	
	Node 7	7→l→f→3→k→h→10→9→6	21,775,187.5	
	Node 8	8→1→j→9→6	8,382,437.5	
	Link f	7→f→3→i→b→c→k→h	18,761,937.5	
	Link g	4→g→i→c→d→m→10	19,844,703.1	
	Link j	j→l→8→e→h→6	3,576,734.4	

5.3.3 Sensitivity of results to multiple parameters

The sensitivity of results to the number of work teams, duration of the planning horizon, and construction cost of projects are analyzed below. Results with modified parameters are compared as shown in Table 8. Sequenced numbers in brackets denote the improvement projects (as labelled in Table 6) to be sequentially completed in the planning horizon.

Table 8 Comparison of optimized results with modified parameters

Number of work teams	Duration of planning horizon (months)	T = 300	T = 300	T = 600
	Construction cost	$C_l'=C_l$, $F'=F$	$C_l'=0.5C_l$, $F'=0.5F$	$C_l'=C_l$, $F'=F$
1 team	Optimal sequence	(1, 2)	(1, 2)	(1, 2, 4)
	Total cost ($\$\times10^8$)	11.2226	9.8521	27.7593
2 teams	Optimal sequence	No action	(1, 2)	(1, 2, 4)
	Total cost ($\$\times10^8$)	8.4450	7.7466	15.0952
3 teams	Optimal sequence	No action	(1)	(1, 2, 4)
	Total cost ($\$\times10^8$)	7.1910	7.1345	14.0312

It is assumed that each specific number of work teams incurs a constant labor cost per month, and this labor cost is not affected by occurrence of disruptions. Therefore, in this case, the excess of cost does not consider the number of work teams. With more teams available, total duration of restoration after occurrence of each disruption event is shortened, which also reduces minimized excess in a disruption scenario. Consequently, the cumulative expected excess decreases, largely contributing to the decrease in total cost. The reduced monthly expected excess also disfavors implementation of long-term improvements because potential savings of cumulative expected excess from completing an improvement may not exceed increases in construction cost.

From Table 8 it is also observed that more improvement projects are completed when the planning horizon increases and the construction costs decrease (although with a proportionally lower budget accumulation rate). Both factors increase the chance that the cumulative saving in monthly expected excess after completing an improvement project exceeds the construction cost of this project.

To examine impacts of long-term demand growth, the following modified numerical case is used. Only one work team is available in a planning horizon of 180 months. Construction costs (C_l) and the budget accumulation rate (F) are halved, with initial demands lower than those previously used for each OD pair. Demand grows exponentially at a constant rate. For each examined sequence of improvements, the planning horizon is divided into sub-periods using completion times as well as multiples of 30 months. Table 9 shows optimized results under multiple demand growth rates and two initial demand levels.

Table 9 Optimized results with varied demand levels

Initial demand 20% below original case				
Annual demand growth rate	0	1.2%	2.4%	3.6%
Optimal improvement sequence	(1)	(1)	(1, 2)	(1, 2)
Total cost ($\$\times 10^8$)	3.6395	4.5920	5.9244	7.2033
Initial demand 10% below original case				
Annual demand growth rate	0	1.2%	2.4%	3.6%
Optimal improvement sequence	(1)	(1, 2)	(4)	(4)
Total cost ($\$\times 10^8$)	5.0067	6.4329	7.9411	8.2637

Results show that a higher initial demand level as well as a higher demand growth rate increase the total cost, without necessarily favoring more improvements in the optimal sequence.

The potential uncertainty of long-term demand growth can be considered using a simplified method. First, for each candidate improvement sequence (and its uniquely determined schedule) its total cost (TC) under each scenario of demand growth is computed. Then, with probabilities of growth scenarios, the expected TC is obtained. These expected values guide the optimization search for the improvement plan. For

example, with an initial demand 10% lower than the original demand and other parameters unchanged, possible cases of exponential annual growth rate of demand are 1.2%, 2.4%, and 3.6%, with probabilities of 20%, 60%, and 20%, respectively. In this example the optimal improvement plan is to sequentially complete projects #1 and #2, with the expected TC of $\$7.8605\times10^8$.

In the original numerical case, the monthly interest rate (r) is zero. If non-zero values of r are applied with other parameters unchanged, the optimized results are compared in Table 10. With a large number of months in the planning horizon, the minimized PV of total cost is rather sensitive to r. A higher r also appears to disfavor implementation of more projects because of the more heavily discounted cumulative savings in expected excess after completing a project.

Table 10 Optimized results with varied interest rates

Monthly interest rate	0	0.1%	0.2%	0.3%	0.4%
Optimal improvement sequence	(1, 2)	(1, 2)	(1, 2)	(1, 2)	(1)
PV of total cost ($\$\times10^8$)	11.2226	10.0302	9.0314	8.1912	7.3862

5.4 Demonstrating the model in a larger network

5.4.1 Numerical case for a larger network

The proposed model and solution methods are further demonstrated in a larger-scale rail freight network shown in Figure 8. The structure of this network is derived from the major freight rail links and stations in the northeastern U.S., while also including Montreal, Canada. Each node represents a city and its metropolitan area. The number labels of these nodes and their corresponding cities are noted below the figure.

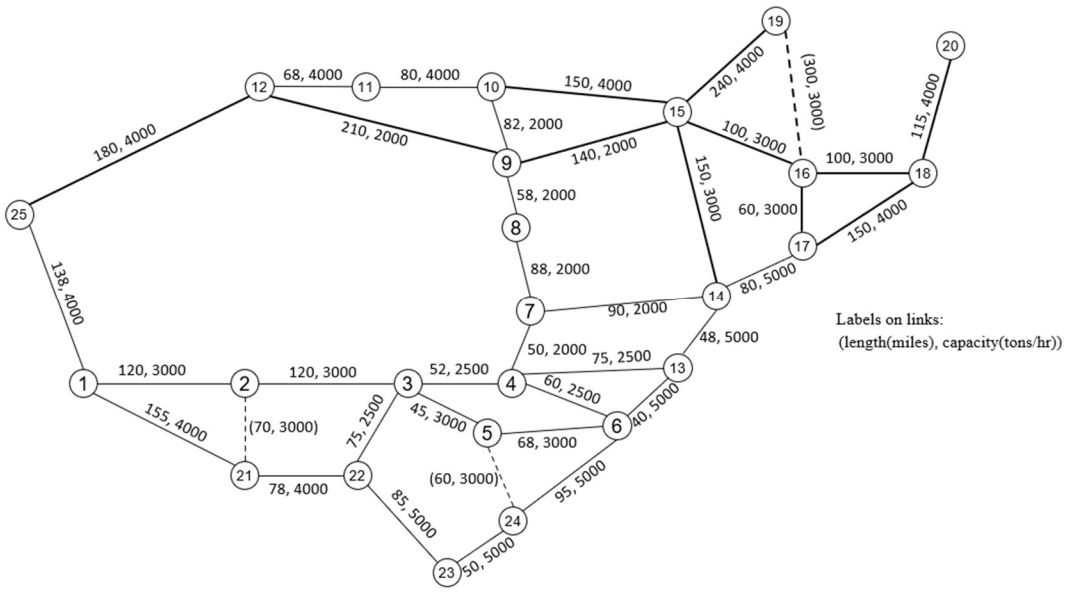


Figure 8 A simplified rail freight network in northeastern US

1- Pittsburgh, PA; 2- Altoona, PA; 3- Harrisburg, PA; 4- Reading, PA; 5- Lancaster, PA; 6- Philadelphia, PA; 7- Allentown, PA; 8- Scranton, PA; 9- Binghamton, NY; 10- Syracuse, NY; 11- Rochester, NY; 12- Buffalo, NY; 13- Trenton, NJ; 14- Newark, NJ/NYC, NY; 15- Albany, NY; 16- Springfield, MA; 17- New Haven, CT; 18- Boston, MA; 19- Montreal, Quebec, CAN; 20- Portland, ME; 21- Cumberland, MD; 22- Hagerstown, MD; 23- Washington, DC; 24- Baltimore, MD; 25- Cleveland, OH

All the links in the network are assumed to be two-directional, and their lengths in miles are approximated from www.distancecalculator.net. The links in dashed lines represent those to be constructed in the possible network improvement projects. One-direction capacity values of the links are synthesized estimates of the real values. The node capacities, are assumed to be 6,000 tons/hr for nodes 5 and 7, 10,000 tons/hr for nodes 2, 3, 4, 8, and 9, 15,000 tons/hr for nodes 19 and 20, and 20,000 tons/hr for the remaining nodes.

A basic gravity model is used for generating the demand matrix that specifies the hourly amounts of freight that must be delivered for each OD pair. To apply that gravity model, the 2021 estimated population of metropolitan areas (as a type of core-based statistical area) provided by census.gov are used as the population values of these nodes. To consider freight demands from outside the network (e.g., from other areas in the North America and major international harbors), extra population numbers are added to the following nodes: 70 million for Node 1 (Pittsburgh), 20 million for Node 6 (Philadelphia), 15 million for Node 12 (Buffalo), 75 million for Node 14 (Newark/NYC), 20 million for Node 18 (Boston), 10 million for Node 19 (Montreal), 10 million for Node 20 (Portland), 70 million for Node 23 (Washington, DC), 20 million for Node 24 (Baltimore), and 90 million for Node 25 (Cleveland). With population and distance values ready, the raw demand value from Node *i* to Node *j* is given by: $(Factor) \cdot \sqrt{Pop_i \cdot Pop_j} \Big/ (Dist_{ij})^{Power}$, where the "Factor" is set to 0.0007 and the "Power" is set to 0.8. In the resulting raw demand values, those below 60 are filtered out and the remaining values are converted to their closest multiples of 25. In the final demand matrix, all demand values are multiples of 25.

For various cost components, the unit costs are c_r=\$1/ton/mi for rail shipment, c_u=\$20/tons/hr for users' time value, and c_{alt}=5c_r=\$5/ton/mi for shipment by an alternative mode. The maximum allowable rail shipment time T_w is assumed to be $2t_w$. Only one work team based at Node 14 is available. Other settings, parameters and assumptions at the lower and middle level model are unchanged from those in the small network case. In the GA, parameters for the short-term restoration optimization are also used here. In the initial population, the probability that each location in a chromosome

is replaced with a blank is set to 0.15. When performing mutation in a chromosome, with a probability of 0.5, a random location is chosen where the number is replaced with a blank (or vice versa).

In this numerical case, there are 6 synthesized disruption scenarios, each having a center of disruption. When a scenario is centered at a node, the node itself, links connecting this node, and outer nodes connected by these links have their capacities reduced by 50%, 30%, and 20%, respectively. If a link becomes the center, then capacities also decrease by 50%, 30%, and 20% for this link, its two end nodes, and for outer links connecting those two nodes, respectively. If the capacity of a damaged link is x (x<1) times of the original, then the travel time through this link is assumed here to be 1/x times of the original. Centers of disruption in these 6 scenarios are Node 3, Node 6, Node 10, and links connecting 3&5, 7&8, and 17&18, with monthly occurrence probabilities (π_ω) of 0.02, 0.04, 0.03, 0.04, 0.05, 0.03, respectively.

10 candidate projects are available for long-term network improvement in a planning horizon of 240 months. With a constant external budget supply of $\$5\times10^5$ per month, details of these projects are shown below.

Table 11 Candidate improvement projects and their parameters in the larger rail network

No.	Action of project	Construction cost C_l(× $\$10^6$)	Budget accumulation time (months)	Minimum work duration (months)
1	Increase capacities of links connecting 4&7, 7&8, 7&14, 8&9 and Node 7 by 50%	14.8	29.6	30
2	Increase capacities of links connecting 1&2, 2&3, 3&5, 5&6 and Node 5 by 1/3	12.1	24.2	24

3	Increase capacities of links connecting 6&24 and 23&24 by 25%	3.625	7.25	12
4	Increase capacities of links connecting 6&13 and 13&14 by 25%	2.2	4.4	6
5	Increase capacities of links connecting 14&17, 17&18, and 18&20 by 25%	8.625	17.25	18
6	Increase capacities of links connecting 9&10, 9&12, and 9&15 by 50%	21.6	43.2	36
7	Increase capacities of links connecting 3&4, 3&22, and 4&6 by 20%	3.74	7.48	12
8	Add a 70-mile link with 3000 tons/hr capacity, connecting 2 and 21	8.4	16.8	15
9	Add a 60-mile link with 3000 tons/hr capacity, connecting 5 and 24	7.2	14.4	15
10	Add a 300-mile link with 3000 tons/hr capacity, connecting 16 and 19	36	72	48

5.4.2 Optimizing long-term improvement of this network

The model and the numerical case are coded in Python 3.7.3. The main program is run on the previously specified computation environment. Nevertheless, this numerical case involves a much larger number of OD pairs and their feasible paths than the smaller numerical case, which greatly increases the computation time for optimizing a lower-level LP model. For a given network state, it takes a CPU core several hours to compute its monthly expected excess value. If the GA is applied to the upper-level for optimizing the selection and sequence of 10 candidate improvement projects, hundreds of network states must be evaluated. If computation is not parallelized among multiple cores, the total computation time for a single core will be unacceptably long, and it will also be rather time-consuming to verify the effectiveness of the proposed GA at the upper level.

To address this problem, the monthly expected excess values of all 1,024 possible network states are computed separately in a high-performance computing (HPC)

environment, namely the Zaratan cluster operated by University of Maryland. The computation is accomplished through two jobs (programs) submitted to the cluster. One job requested 64 CPU cores (AMD EPYC 7763 @2.45 GHz base and 3.5 GHz turbo) to compute the excess value of 128 network states in parallel (each core being assigned two states to evaluate), and the other job requests 128 cores to compute the excess value of the remaining 896 network states (each core being assigned seven states to evaluate). The first job takes 11:00:48 to complete and the second job takes 42:55:04.

After all the monthly expected excess values are ready, they are imported to the personal laptop and the upper-level optimization using GA can be easily performed and verified. 5 runs of the upper-level model GA take an average of 0.301 seconds, and they return the same optimized selection and sequence: projects 3, 7, and 4 are selected and sequentially completed within the planning horizon. With budget accumulation times and minimum work durations of these projects, the schedule is also uniquely determined: projects 3, 7, and 4 are completed at 12, 19.25, and 20.73 months into the planning horizon, respectively. The resulting objective function value (accumulated expected excess plus project construction cost) is $\$3.4538\times10^8$. If both the construction cost and the external budget shrink by 25%, then the optimized selection and sequence is to sequentially complete projects 3, 7, 4, and 9 at 12, 19.25, 20.73, and 34.13 months into the planning horizon, respectively.

5.4.3 Validating quality of the GA-generated solution

The quality of this GA-optimized result is statistically checked. In this numerical case, the total number of all possible selections and sequences of improvement projects is:

$$\sum_{i=0}^{10} \frac{10!}{i!} = 9{,}864{,}101$$

From all these possible plans, 100,000 of them (~1%) are randomly sampled. The distribution of objective function values corresponding to these samples are shown in the left histogram below. It appears to resemble that of the Gumbel distribution (skewed to the left). Using the "stats" module of the Python package "SciPy", the parameters for the best-fitting Gumbel distribution are found (loc = 483,441,820, scale = 18,637,600). The curve of the resulting probability density function appears to fit well to the normalized histogram, as shown below to the right. The Kolmogorov-Smirnov test is performed for the goodness of fit. With a sample size of 1,000, the resulting p-value is 0.642, which means that the null hypothesis – the sample is distributed according to the Gumbel distribution – cannot be rejected at a confidence level of 95%.

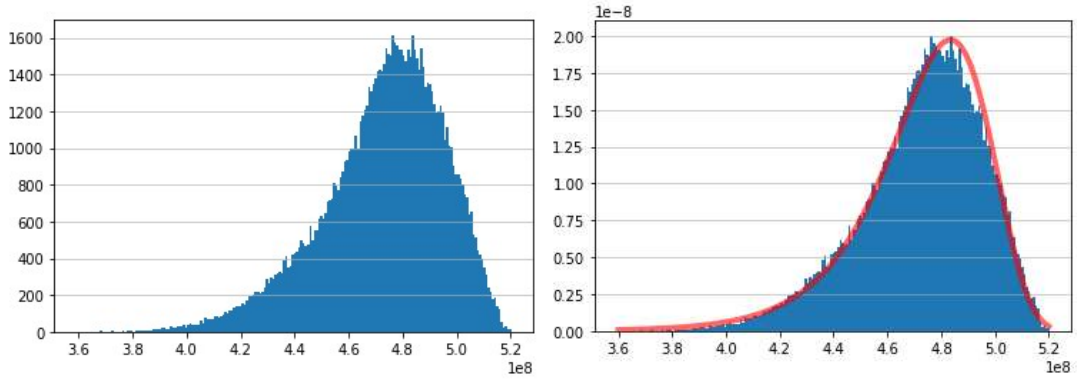


Figure 9 Histograms of OF values from 100,000 sampled improvement plans

The GA-minimized OF value ($3.4538×10^8) corresponds to the cumulative density function value of 0.0006 in the best-fitting Gumbel distribution, which indicates that this minimized value is within the smallest 0.1% of the population and shows high quality of the GA result. The GA-minimized value is even smaller than the smallest OF value in the 100,000 sampled improvement plans ($3.5908×10^8).

Chapter 6: Numerical results for Problem 2

6.1 The numerical case

6.1.1 The example network, the candidate projects, and their parameters

To demonstrate the proposed integrated optimization model for Problem 2, a road network based on the shape and topology of the Sioux Falls network (LeBlanc et al., 1975) is used in the numerical case, as shown in Figure 10.

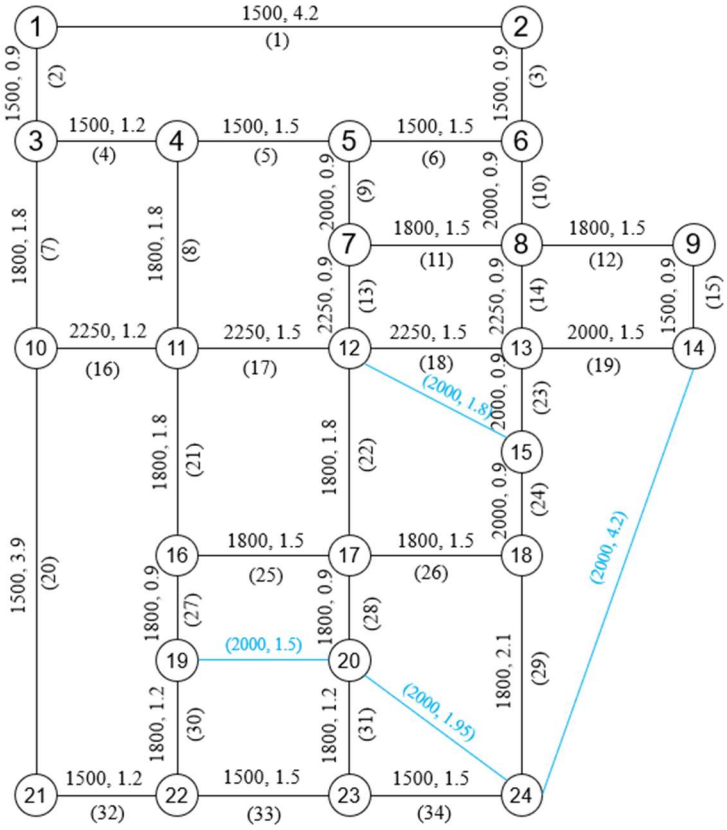


Figure 10 The example road network

All links are bidirectional. Numbers above (at the left of) each link are its one-directional normal capacity in vehicles/hour and its length in miles. Each existing link has its numeric label. The blue links represent links that are currently unavailable but may be added in candidate improvement projects. This network, with its components

and their attributes, are modeled using NetworkX, a Python package. This package also assists the Frank-Wolfe algorithm by finding the shortest path for each OD pair.

Original hourly demands in veh/hr by OD pair during peak and off-peak hours are shown in Table 12 and Table 13, respectively:

Table 12 Original hourly demands by OD pair during peak hours

To \ From	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1		180							270				240				150							
2							240			270						120						90		
3					210	150					210			210										
4								180				180						120			180			
5			150							210					180				90					
6	240							240									180						150	
7				240					300				300											240
8		300									330				300					240				
9						300						360						240			270			
10					330		330								270				240					
11		240										270								270		240		
12								330									270	270					270	
13			360							330						240					300			
14	270						300												300					210
15				270							300			300								240		
16	210				180				240											150				
17						210					240					180								120
18					210									240							150		210	
19			180				210										120					180		
20				150								150		180									120	
21		160							300				210					180						
22			210					240							150									120
23	150					180							240						150					
24				180						240						150				180				

Table 13 Original hourly demands by OD pair during off-peak hours

To \ From	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1		90							120				90				90							
2							120			150						30						60		
3					120	90					120			90										
4								90				120						60			60			
5			60							120					120				30					
6	120							120									90						60	
7				120					180				120											120
8		150									210				150					120				
9						150						180						120			120			
10					180		150								120				150					
11		120										120								120		120		
12								120									150	150					150	

13			180							210						120					180			
14	150						150												180					90
15				150							180			210								120		
16	120				90				90											90				
17						120					150					90								60
18					120									120							90		120	
19			90				90										60					60		
20				60								60		90									60	
21		60							180				120					90						
22			90					150							60									60
23	60					60							90						60					
24				120						150						60				90				

The peak hour takes up 25% of all time (α=0.25), and the rest is the off-peak hour.

The 10 candidate improvement projects are listed in Table 14, with their base values of construction cost (C_l) and required work time (ϕ_l).

Table 14 Candidate improvement projects and their parameters in the road network

Proj. #	Description	Construction cost (×\$$10^7$)	Required work time (years)
1	Increasing capacities of links 4, 5, 6 by 1/3	2.8	0.75
2	Increasing capacities of links 11, 12, 15 by 1/3	2.6	0.75
3	Increasing capacities of links 25, 26, 29 by 25%	3.4	1
4	Increasing capacities of links 8, 21, 27, 30 by 25%	2.85	1
5	Increasing capacities of links 22, 28, 31 by 25%	1.95	0.75
6	Increasing capacities of links 20, 32, 33, 34 by 20%	5.4	1.5
7	Adding a link connecting 12 and 15	3.6	1.25
8	Adding a link connecting 19 and 20	3.0	1
9	Adding a link connecting 20 and 24	3.9	1.25
10	Adding a link connecting 14 and 24	8.4	2

When projects 1, 2, 3, 6 are under construction, capacities of affected links decrease by 25%, and travel times through them increase by 1/3. When projects 4 and 5 are under construction, capacities of affected links decrease by 20%, and travel times through them increase by 25%.

In this road network, the free-flow travel time through each link in normal operation (t_a^0) is determined by the free-flow speed of 30 mph. The unit value (cost) of travel time (v) is \$15/veh/hr. The planning horizon lasts for T = 15 years, with H = 8,760 equivalent hours per year. This planning horizon is first divided into 30 sub-periods of 0.5 years, and then further divided by the starting and completion times of projects. The annual interest rate r is 5%. A fraction ρ = 1% of total travel time cost is used as the internal source of the improvement budget. In the F-W traffic assignment algorithm the threshold of Δ_x in equation (37) is 0.001. When determining the budget-ready times of implemented projects, the threshold of Δ in equation (34) is \$1,000.

6.1.2 Correlated uncertainties and their parameters in multiple scenarios

In this numerical case, the three correlated uncertain parameters: the annual growth rate of demand (g), the annual external budget supply (F), and the multiplier of required construction time of projects (ψ), are assumed to follow the multivariate normal distribution as given by:

$$(g, F, \psi)^T \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{16'}$$

where $\boldsymbol{\mu}$ is the mean vector, and $\boldsymbol{\Sigma}$ is the covariance matrix:

$$\boldsymbol{\mu} = \left(\mu_g, \mu_F, \mu_\psi\right)^T \tag{43}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_g^2 & \rho_{gF}\sigma_g\sigma_F & \rho_{g\psi}\sigma_g\sigma_\psi \\ \rho_{gF}\sigma_g\sigma_F & \sigma_F^2 & \rho_{F\psi}\sigma_F\sigma_\psi \\ \rho_{g\psi}\sigma_g\sigma_\psi & \rho_{F\psi}\sigma_F\sigma_\psi & \sigma_\psi^2 \end{bmatrix} \tag{44}$$

Here, the means, standard deviations, and correlation coefficients are:

$$\mu_g = 0.025 = 2.5\%, \mu_F = \$1.5 \times 10^7, \mu_\psi = 1$$

$$\sigma_g = 0.0025, \sigma_F = \$1.0 \times 10^6, \sigma_\psi = 0.1$$

$$\rho_{gF} = 0.6,\ \rho_{g\psi} = -0.2, \rho_{F\psi} = 0.3$$

From the multivariate normal distribution specified above, 50 vectors of uncertain parameters are sampled, each vector representing a scenario $\omega \in \Omega$. The sampling process has three steps:

First, 50 3-dimensional vectors are sampled in unit cube where each dimension has an independent uniform distribution $U(0,1)$, using a quasi-Monte Carlo sampling method: Hammersley Sampling (Kalagnanam and Diwekar, 1997). This technique generates a series of sample points with a relatively small sample size but a high uniformity (namely, low discrepancy) in the sampling space. Based on the Hammersley sequence, the m^{th} (m = 1, 2, …, 50) sample point $\boldsymbol{x}_m$ is given by the following rules:

1) The first entry of the vector is $x_{m1} = (m - 0.5)/50$.
2) Write the integer m as $m = \sum_{i=0}^{\lfloor \log_2 m \rfloor} m_i \cdot 2^i$. That is, use a radix of 2 to represent m. The second entry is given by $x_{m2} = \sum_{i=0}^{\lfloor \log_2 m \rfloor} m_i \cdot 2^{-i-1}$.
3) Similarly, write m as $m = \sum_{i=0}^{\lfloor \log_3 m \rfloor} m_i \cdot 3^i$. The third entry is $x_{m3} = \sum_{i=0}^{\lfloor \log_3 m \rfloor} m_i \cdot 3^{-i-1}$.

Second, for each entry of each sample vector, use the inversed cumulative density function of the standard normal distribution $N(0,1)$ to map the entry value to its corresponding percentile in $N(0,1)$. For example, 0.5 is mapped to the 50$^{\text{th}}$ percentile which is 0, and 0.9 is mapped to the 90$^{\text{th}}$ percentile of 1.28. This results in 50 sample points $\boldsymbol{x}'_m$ in a multivariate standard normal distribution with 3 independent dimensions.

Third, using Cholesky decomposition, a unique lower-triangle matrix $\boldsymbol{A}$ is found such that $\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{\Sigma}$. For each sample point $\boldsymbol{x}'_m$, calculate the vector $\boldsymbol{x}''_m = \boldsymbol{\mu} + \boldsymbol{A}\boldsymbol{x}'_m$.

According to the affine transformation property, the resulting set of 50 sample points now follows the desired multivariate normal distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

The three uncertain parameters in the 50 scenarios are listed in Table 15:

Table 15 Values of uncertain parameters in sample scenarios

Scen. #	g(%)	F($10^7)	ψ	Scen. #	g(%)	F($10^7)	ψ	Scen. #	g(%)	F($10^7)	ψ
1	1.918	1.360	1.011	18	2.404	1.431	0.858	35	2.624	1.588	1.135
2	2.030	1.333	1.038	19	2.417	1.542	1.028	36	2.638	1.447	0.796
3	2.089	1.455	0.967	20	2.430	1.402	1.006	37	2.653	1.566	0.982
4	2.131	1.319	0.958	21	2.443	1.519	0.952	38	2.669	1.518	1.019
5	2.165	1.445	1.107	22	2.456	1.470	0.995	39	2.685	1.643	0.968
6	2.193	1.401	0.945	23	2.469	1.598	1.158	40	2.702	1.435	0.908
7	2.218	1.524	1.094	24	2.481	1.390	0.888	41	2.719	1.568	1.071
8	2.241	1.315	1.041	25	2.494	1.517	1.040	42	2.739	1.522	0.907
9	2.261	1.455	0.880	26	2.506	1.469	1.126	43	2.759	1.638	1.051
10	2.281	1.408	0.965	27	2.519	1.585	0.866	44	2.782	1.501	1.061
11	2.298	1.523	1.107	28	2.531	1.445	0.924	45	2.807	1.616	0.891
12	2.315	1.385	0.882	29	2.544	1.557	1.065	46	2.835	1.571	0.950
13	2.331	1.499	1.035	30	2.557	1.507	0.896	47	2.869	1.723	1.115
14	2.347	1.451	1.078	31	2.570	1.666	1.083	48	2.911	1.465	0.809
15	2.362	1.590	1.038	32	2.583	1.348	0.947	49	2.970	1.622	0.975
16	2.376	1.321	0.931	33	2.596	1.526	0.934	50	3.082	1.597	1.016
17	2.390	1.480	1.133	34	2.610	1.476	0.972				

6.2 Optimized long-term improvement of the network

6.2.1 The optimized improvement plan with detailed results in all scenarios

The model and the numerical case are coded in Python 3.7.3. Under the personal computation environment as specified in Section 5.1, given an improvement sequence σ_{imp} and a scenario ω, computing the PV of cumulative travel time cost (given by equation (18)) plus implementation cost takes 1 to 2 minutes. With hundreds of different improvement sequences, each encompassing 50 scenarios, to be evaluated in the upper-level GA, loading a single CPU core with all the tasks will lead to hundreds

of hours of computation time which is undesirable for practical use. Therefore, the Zaratan cluster (as mentioned in 5.5.2) is used for parallelizing the computation. 51 CPU cores are used, among which one core acts as the "leader" that performs the GA iterations. After converting each chromosome into an improvement sequence, each sequence that has not been previously evaluated in the current GA search is dispatched from the "leader" to the remaining 50 "follower" cores. The "followers" No.1 to No.50 are assigned with the scenarios No.1 to No.50, respectively. Each "follower" computes the PVC under the assigned scenario and the improvement sequence it receives from the "leader". After the computation is finished, each "follower" reports its result to the "leader", who computes the objective function value of the newly evaluated sequence (as in equation (19)) as soon as it receives responses from all the "followers". As the above process repeats, the "leader" obtains the fitness values of chromosomes and conducts the operations and iterations of GA until the algorithm terminates.

The GA parameters and operator settings are the same as those used in the upper-level improvement optimization in Problem 1. With the same GA settings and the same structure and shape of chromosomes, the performance of GA in optimizing the improvement plan in the case of Problem 2 is equally good as that in the larger network case of Problem 1. After a parallel computation time of 23:57:35, the optimized selection and sequence of improvement projects is obtained. The solution is represented by (7, 8, 9, 10, 1), which means that projects 7, 8, 9, 10, and 1 are selected and sequentially completed within the planning horizon. The resulting minimized expected PVC is \$4.67761 $\times 10^9$. Under this optimized improvement sequence, the PVC and the

completion time of the last project in the sequence (last completion time, LCT) in each scenario are listed in Table 16:

Table 16 Results under the optimized improvement sequence across sample scenarios

Scen. #	PVC($\$10^9$)	LCT(year)	Scen. #	PVC($\$10^9$)	LCT(year)	Scen. #	PVC($\$10^9$)	LCT(year)
1	4.45669	6.2140	18	4.63668	5.9092	35	4.72813	5.8026
2	4.49955	6.3047	19	4.64240	5.9563	36	4.72895	5.6313
3	4.51672	5.8909	20	4.65076	5.6697	37	4.73505	5.8938
4	4.53672	6.3499	21	4.65143	6.0451	38	4.74344	5.5709
5	4.54909	5.9928	22	4.65871	5.6837	39	4.74569	5.7163
6	4.55767	6.0625	23	4.66679	5.8306	40	4.75690	5.3784
7	4.56733	5.7697	24	4.66914	5.6287	41	4.76441	5.9279
8	4.58050	6.3585	25	4.67366	6.0852	42	4.76941	5.6278
9	4.58092	5.8839	26	4.68211	5.7394	43	4.77885	5.6632
10	4.59133	6.0338	27	4.67800	5.9290	44	4.79127	5.4490
11	4.59878	5.7814	28	4.68811	5.4925	45	4.79453	5.7880
12	4.60398	6.1116	29	4.69348	5.9023	46	4.80881	5.3987
13	4.61011	5.7896	30	4.69595	5.6570	47	4.82625	5.5291
14	4.61843	5.9498	31	4.70302	5.7126	48	4.84107	5.3093
15	4.61986	5.5628	32	4.71236	5.4151	49	4.86484	5.8268
16	4.63075	6.3298	33	4.71199	6.2246	50	4.91398	5.4230
17	4.63631	6.2140	34	4.71963	5.6553			

With the optimized improvement sequence, the standard deviation of PVC across all scenarios is $\$9.65775 \times 10^7$. The average LCT is 5.8074 years into the planning horizon, and the standard deviation is 0.2675 years. The coefficient of variation (CV), expressed as the ratio of the standard deviation to the mean of a variable, is 0.02065 for PVC and 0.04606 for LCT. In comparison, the CVs of the uncertain parameters g, F, and ψ across the 50 scenarios are 0.09974, 0.06472, and 0.08947, respectively.

To show more details of underlying computation results, scenario No. 20 is taken as an example. With g=2.430%, F=$\$1.402 \times 10^7$, ψ=1.006, and the improvement sequence of (7, 8, 9, 10, 1), the starting time, budget ready time, and completion time

of each implemented project in the planning horizon are listed in Table 17. The starting times of projects 8, 9, and 10 are earlier than the completion times of projects 7, 8, and 9, respectively, meaning that the constructions of projects 8, 9, and 10 overlap with those of projects 7, 8, and 9, respectively.

Table 17 Starting and completion times of implemented projects in a scenario

Proj. #	7	8	9	10	1
Proj. start time (years)	0	1.0125	1.8545	2.9457	5.2755
Budget ready time (years)	1.0125	1.8545	2.9457	5.2755	6.0451
Proj. completion time (years)	1.2575	2.0185	3.1119	5.2755	6.0451

As previously mentioned, the planning horizon is divided into sub-periods by multiples of 0.5 years and the starting and completion times of projects. With the same set of dividing times, the change in average hourly travel time cost (Z_{UE}, not discounted to PV) with the optimized improvement plan is compared with that without any network improvement in the first 8 years of the planning horizon, as shown in Figure 11:

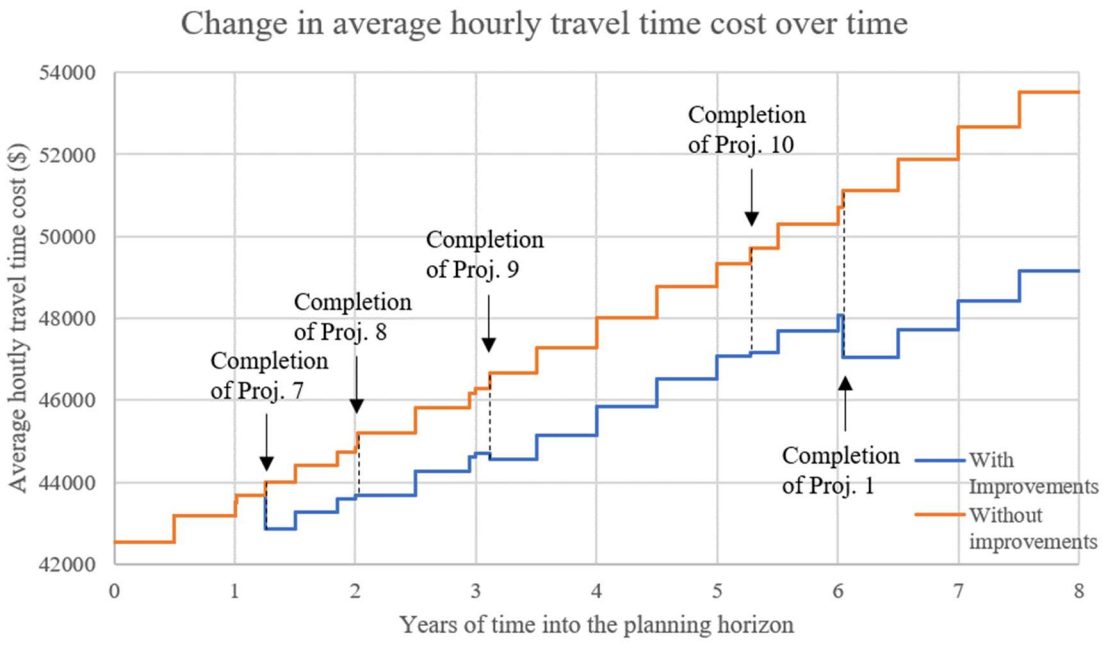


Figure 11 Effects of improvements in average hourly travel time cost

The staircase-shaped graphs stem from the approximation method as specified in Section 4.1.2. It is explicitly shown that the completion of each project reduces the average hourly travel time cost in all later sub-periods and widens the gap between these two graphs. Project 1 is the last to be implemented in the sequence, but it appears to cut off the most future value of Z_{UE}. The reduction in Z_{UE} by completing Project 10 is partially offset by the start of Project 1 which temporarily decreases capacity of and increases travel time through some affected links.

The fraction ρ of total travel time cost used for internal budget supply is 0.01 in the base numerical case. If larger values of ρ are used, the resulting improvement schedules with the sequence (7, 8, 9, 10, 1) and the scenario No. 20 are listed in Table 18.

Table 18 Starting and completion times of implemented projects in a scenario with larger ρ

ρ = 0.02					
Project #	7	8	9	10	1
Project start time (years)	0	0.8367	1.5309	2.4323	4.3528
Budget ready time (years)	0.8367	1.5309	2.4323	4.3528	4.9850
Project completion time (years)	1.2575	1.8427	2.7883	4.4442	5.1073
ρ = 0.03					
Project #	7	8	9	10	1
Project start time (years)	0	0.7132	1.3030	2.0721	3.7108
Budget ready time (years)	0.7132	1.3030	2.0721	3.7108	4.2462
Project completion time (years)	1.2574	1.7192	2.5604	4.0840	4.4653
ρ = 0.04					
Project #	7	8	9	10	1
Project start time (years)	0	0.6217	1.1352	1.8046	3.2359
Budget ready time (years)	0.6217	1.1352	1.8046	3.2359	3.7012
Project completion time (years)	1.2575	1.6277	2.3926	3.8166	3.9904

It is observed that a slightly higher ρ significantly advances the budget ready time of each project, leading to an earlier starting time of its succeeding project and a larger

overlapping part of implementation durations of two sequentially adjacent projects. As a result, the completion times of projects are also advanced, which starts the savings in travel time cost from earlier times and decreases the PVC: when ρ is 0.01, 0.02, 0.03, and 0.04, the PVC in $\$10^9$ is 4.65076, 4.64272, 4.63933, and 4.63686, respectively.

6.2.2 The optimized improvement plan with buses

In the numerical case with the mode of bus added, there are three bus routes in the network: Route 1 runs through nodes 2, 6, 8, 13, 15, 18, and 24 and links 3, 10, 14, 23, 24, and 29. Route 2 runs through nodes 4, 11, 16, 19, and 22 and links 8, 21, 27, and 30. Route 3 runs through nodes 10, 11, 12, 13, and 14 and links 16, 17, 18, and 19. With these routes, travelers of the following 7 OD pairs may use the bus: 2-8, 8-6, and 15-8 served by Route 1, 22-11 and 22-19 served by Route 2, and 10-13 and 12-11 served by Route 3.

The passenger capacity of a bus (K_{bus}) is 15, and the average number of people in a car (K_{car}) is 1.25. When considering two modes, the unit of original hourly demands by OD pair during peak and off-peak hours is in persons/hr instead of vehs/hr, so the values of q_w in Table 12 and 13 are multiplied by 1.25 while the unit value of travelers' time in dollar per person hour is v=15/1.25=12. Each bus is equivalent to b=2 cars when computing traffic flows and travel times. For the bus passengers, a maximum wait time ($t_{maxwait}$) of 0.25 hours is set to prevent infinite wait time, and the unit value of wait time is v_{wait}=1.5v=18 \$/person hour. The bus fare (c_{bus}) is \$1.2 per passenger, and the unit fuel cost by car (c_{fuel}) is \$0.06 per vehicle mile.

At the beginning of the iteration of mode shares, the initial bus shares for the 7 OD pairs mentioned above are 0.5. In this numerical case, the input initial mode shares are

iterated 4 times with the method specified in Section 4.4, and the UE traffic flows resulting from the mode shares after 3 iterations are used for computing the hourly total travel time cost, whose value is used in the upper-level computation. Such iterations approximately quadruple the computation time spent in F-W traffic assignment. To compress the time needed to obtain the optimized improvement sequence, the following method is applied for parallelizing the computation. 201 CPU cores are used, among which one core acts as the "leader" that performs the GA iterations, dispatches tasks to the "followers" and collects results from the "followers". The 200 "follower" cores are divided into 4 groups, each consisting of 50 cores, one of which is the "sub-leader" of this group. At the beginning of each iteration of GA, the chromosomes are converted to improvement sequences, and each of the first 4 "new" sequences that have not been previously evaluated in the current GA search are sent by the "leader" to each group of "follower" cores. Within each group, each core is assigned with each of the 50 scenarios and computes the PVC under the assigned scenario and improvement sequence. After the computation is finished, each member in the group (including the "sub-leader" itself) reports its result to the "sub-leader" who, upon receiving responses from all group members, computes the objective function value (expected PVC) of the newly evaluated sequence and reports it to the "leader". As long as there are "new" improvement sequences in the current iteration of GA, the "leader" immediately sends the next "new" sequence to the group whose "sub-leader" has just reported its result. The four groups are kept busy until all "new" sequences have been distributed and there are 3 or fewer remaining sequences with pending results, which leaves some groups temporarily idle. When all the "new" sequences in the current iteration have been

evaluated, the "leader" performs the GA operators and starts the next iteration. The GA iteration process loops until the termination condition is reached.

The integrated model with mode choice is demonstrated in numerical cases with and without dedicated bus lanes. When all buses are operated in dedicated bus lanes, the bus lane capacities in vehicles/hour for routes 1, 2, and 3 are 600, 600, and 750, respectively. These capacity values are not affected by improvement projects. For every link that is occupied by a bus route, the route's bus lane capacity is deducted from the link's full capacity, and the remaining capacity is available to cars. When all buses share roads with cars, however, each link on a bus route has its full capacity available to a mixed traffic of bus and car. All other parameters in the numerical cases are unchanged from those specified in Section 6.1.

Under two different modes of bus operation, the optimized improvement sequence, the resulting expected PVC and average LCT with their change rates from the base numerical case, and the standard deviations and CVs of PVC and LCT across 50 scenarios are shown in Table 19.

Table 19 Optimized results with bus and car

Bus operation mode	Dedicated	Mixed with cars
Optimized improvement sequence	(7, 8, 10, 9, 1, 5)	(7, 8, 9, 10, 1, 5)
Expected PVC	$\$4.70346 \times 10^9$	$\$4.69232 \times 10^9$
Change rate from the base case	0.553%	0.314%
Standard deviation of PVC	$\$9.75528 \times 10^7$	$\$9.77401 \times 10^7$
CV of PVC	0.02074	0.02083
Average LCT	6.5066 years	6.5101 years
Change rate from the base case	12.04%	12.10%
Standard deviation of LCT	0.2695 years	0.2698 years
CV of LCT	0.04142	0.04144

In both cases the optimized improvement sequence has 6 projects to be implemented, among which Project 5 is the additional one compared to the result in the base case. The minimized expected PVC is slightly higher than that in the base case, most attributable to the additional wait time spent by bus passengers with a higher unit time cost than the in-vehicle time. The expected PVC with dedicated lanes is higher than that with mixed traffic because the reduction in link capacity available for cars on bus routes increases their travel time, while using dedicated lanes does not save much travel time for buses if the mixed traffic flows are far from congested. The CVs of PVC are both at a similar level to that in the base case (0.02065). Despite the average LCT in both cases being 12% later than in the base case, the standard deviations of LCT are less than 1% larger than that in the base case, which results in noticeably smaller CVs of LCT.

To show more details of underlying mode choice results at the lower-level model, scenario No. 20 is taken as an example. For the 7 OD pairs served by bus routes, the changes in bus shares after each iteration of mode choice during peak and off-peak hours in the first and the last sub-period of the planning horizon are shown in the following tables. The results with dedicated bus lanes and mixed traffic under the corresponding optimized improvement sequence are listed in Table 20 and Table 21, respectively.

Table 20 Iterated bus shares with dedicated bus lanes

In the first sub-period (start of planning horizon)							
OD pair	2-8	8-6	10-13	12-11	15-8	22-11	22-19
Peak – After Iteration 1	0.1999	0.1508	0.3351	0.1886	0.2026	0.3138	0.1741
Peak – After Iteration 2	0.0851	0.0594	0.2768	0.1475	0.0846	0.2289	0.1167
Peak – After Iteration 3	0.0083	0.0055	0.2403	0.1238	0.0081	0.1692	0.0816

Peak – After Iteration 4	0.0062	0.0041	0.2111	0.1055	0.0061	0.1078	0.0499
Off-peak – After Iteration 1	0.0886	0.0741	0.1767	0.1185	0.0881	0.1351	0.0920
Off-peak – After Iteration 2	0.0047	0.0037	0.0339	0.0206	0.0046	0.0059	0.0038
Off-peak – After Iteration 3	0.0047	0.0038	0.0072	0.0042	0.0046	0.0060	0.0038
Off-peak – After Iteration 4	0.0047	0.0038	0.0072	0.0042	0.0046	0.0060	0.0038
In the last sub-period (end of planning horizon)							
OD pair	2-8	8-6	10-13	12-11	15-8	22-11	22-19
Peak – After Iteration 1	0.2687	0.1902	0.4179	0.2443	0.2265	0.3647	0.2014
Peak – After Iteration 2	0.2015	0.1339	0.4052	0.2335	0.1659	0.3280	0.1688
Peak – After Iteration 3	0.1570	0.1019	0.4014	0.2318	0.1275	0.3132	0.1585
Peak – After Iteration 4	0.1161	0.0738	0.4022	0.2304	0.0936	0.3066	0.1535
Off-peak – After Iteration 1	0.1375	0.1112	0.2427	0.1612	0.1249	0.1886	0.1272
Off-peak – After Iteration 2	0.0114	0.0087	0.1361	0.0835	0.0101	0.0449	0.0275
Off-peak – After Iteration 3	0.0051	0.0039	0.0477	0.0277	0.0045	0.0065	0.0039
Off-peak – After Iteration 4	0.0051	0.0039	0.0079	0.0045	0.0045	0.0066	0.0039

Table 21 Iterated bus shares with mixed traffic

In the first sub-period (start of planning horizon)							
OD pair	2-8	8-6	10-13	12-11	15-8	22-11	22-19
Peak – After Iteration 1	0.1380	0.1338	0.1661	0.1441	0.1383	0.1780	0.1602
Peak – After Iteration 2	0.0242	0.0231	0.0488	0.0431	0.0244	0.0752	0.0666
Peak – After Iteration 3	0.0037	0.0035	0.0042	0.0036	0.0036	0.0101	0.0089
Peak – After Iteration 4	0.0037	0.0035	0.0042	0.0036	0.0037	0.0040	0.0035
Off-peak – After Iteration 1	0.0728	0.0700	0.1176	0.1046	0.0728	0.0991	0.0881
Off-peak – After Iteration 2	0.0036	0.0035	0.0047	0.0042	0.0036	0.0040	0.0035
Off-peak – After Iteration 3	0.0036	0.0035	0.0041	0.0036	0.0036	0.0040	0.0035
Off-peak – After Iteration 4	0.0036	0.0035	0.0041	0.0036	0.0036	0.0040	0.0035
In the last sub-period (end of planning horizon)							
OD pair	2-8	8-6	10-13	12-11	15-8	22-11	22-19
Peak – After Iteration 1	0.1651	0.1598	0.1902	0.1702	0.1665	0.2023	0.1826
Peak – After Iteration 2	0.0666	0.0645	0.0991	0.0885	0.0659	0.1238	0.1120
Peak – After Iteration 3	0.0073	0.0071	0.0352	0.0314	0.0073	0.0729	0.0649
Peak – After Iteration 4	0.0036	0.0035	0.0040	0.0036	0.0036	0.0265	0.0233
Off-peak – After Iteration 1	0.1070	0.1029	0.1525	0.1358	0.1071	0.1354	0.1210
Off-peak – After Iteration 2	0.0037	0.0035	0.0335	0.0295	0.0037	0.0143	0.0125
Off-peak – After Iteration 3	0.0036	0.0035	0.0040	0.0036	0.0036	0.0040	0.0035
Off-peak – After Iteration 4	0.0036	0.0035	0.0041	0.0036	0.0036	0.0040	0.0035

It is noticeable from the tables above that in many situations the bus shares quickly converge to nearly zero for these OD pairs. During off-peak hours such convergence always occurs on all these OD pairs and is much faster than in peak hours. This phenomenon stems from the low bus demand. When the initial demand of bus passengers (with an initial bus share of 50% for each OD pair) is very low for a route, there will be a very small number of buses operating on the route, which causes a long wait time for each bus passenger. As a result, the travel impedance by bus is much higher than that by car, encouraging most travelers to drive. This further decreases the bus demand and, eventually, significantly increases the travel impedance by bus, while the increase in the impedance by car due to higher traffic flows is much smaller. Note that the bus shares above do not actually converge to zero because the constraint on maximum wait time for bus prevents the bus impedance from going to infinity. At the end of the planning horizon when the base demand values are much higher than at the beginning, the iteration of bus shares towards zero is slower.

Given the same initial demand of bus passengers, each passenger's wait time under the two different modes of bus operation will be the same. In this numerical case, if buses use dedicated lanes, a passenger's in-vehicle time will be close to the free-flow travel time, while cars traveling beside the bus lanes will experience increased travel time due to reduced capacities. If buses share roads with cars, a passenger will experience longer in-vehicle time than on dedicated lanes, whereas car users will travel faster on bus routes with full road capacity available. As a result, the advantage to travel by car is more pronounced in the mixed traffic case. The bus shares of all 7 OD pairs converge towards zero with mixed traffic even in peak hours, while with dedicated bus

lanes some OD pairs appear to have their bus share converging at values far above zero during peak hours. With much higher demands at the end of the planning horizon, more OD pairs have their bus shares converging at higher values. The observations above reveal that the implementation of bus service is more desirable when buses are to be operated on dedicated lanes and there is a sufficiently high potential travel demand along bus routes.

6.3 Sensitivity analysis

6.3.1 Sensitivity to demand level, users' value of time, and construction cost

The sensitivity of results to demand level (q_w), unit value of users' travel time (v), and construction cost of projects (C_l) are analyzed below. The results to be compared include the improvement sequence optimized by GA, the mean (expectation) and the standard deviation of PVC and LCT across 50 scenarios, and the CV of PVC and LCT. The change rates of the average PVC and LCT in comparison to those in the base case are calculated.

By increasing or decreasing all numbers in the peak and off-peak demand matrices (Table 12 and Table 13) by the same percentage and keeping all other parameters in the base numerical case unchanged, four modified cases with different demand levels are generated and the results obtained under these cases are compared in Table 22. It is observed that a higher demand level encourages the implementation of more improvement projects in the optimized sequence. With congestion effects in the network (see the BPR function in equation (9)), a higher demand level increases the cumulated savings in average hourly travel time cost by completing a project. Due to the congestion effects, the minimized expected PVC sees larger change rates than the

demand level and a higher elasticity with respect to demand when demand is higher, with the CV of PVC also showing an upward trend. The average LCT is largely influenced by the number of projects to be implemented in the optimized sequence. Under the same optimized sequence, a higher demand level increases the rate of collecting a small fraction of total travel time cost as a source of budget supply, thereby shortening the time needed to complete all projects. For the same reason, the completion time of each project is more likely to be decided by its required work time (ϕ_l) in a scenario, which explains why the standard deviation and the CV of LCT also decrease.

Table 22 Optimized results with changes in the demand level

Demand change	-20%	-10%	10%	20%
Optimized sequence	(7, 10, 8, 9)	(7, 10, 8, 9)	(7, 8, 9, 10, 1, 5)	(7, 8, 10, 9, 1, 5)
Expected PVC ($\$10^9$)	\$3.57546	\$4.11535	\$5.26306	\$5.87338
Change rate from base	-23.56%	-12.02%	12.52%	25.26%
Std. dev. of PVC($\$10^7$)	\$7.13862	\$8.41062	\$11.01443	\$12.55107
CV of PVC	0.01997	0.02044	0.02093	0.02137
Average LCT (year)	5.4545	5.3476	6.3637	6.2186
Change rate from base	-6.08%	-7.92%	9.58%	7.08%
Std. dev. of LCT (year)	0.2225	0.2114	0.2707	0.2574
CV of LCT	0.04079	0.03954	0.04254	0.04139

Another four modified cases are generated by only changing the value of users' travel time (v) from the base case. The optimization results under these cases are compared in Table 23. The increase in v also increases the cumulated savings in travel time cost by completing an improvement project and therefore encourages the implementation of more projects in the planning horizon. The expected PVC has an approximately linear increase with the higher v, while its change rate is slightly lower than that of v

because the PV of construction cost, as a small part of the PVC, is not directly affected by v. The CV of PVC nearly stays constant and does not show a clear trend with changes in v. Under the same optimized sequence, a higher v affects the average, the standard deviation, and the CV of LCT in the same manner as a higher demand level does, for the same reasons.

Table 23 Optimized results with changes in the value of travel time

Value of time change	-20%	-10%	10%	20%
Optimized sequence	(7, 10, 8, 9)	(7, 8, 9, 10, 1)	(7, 8, 9, 10, 1, 5)	(7, 8, 9, 10, 1, 5)
Expected PVC ($\$10^9$)	\$3.76175	\$4.21990	\$5.13499	\$5.59158
Change rate from base	-19.58%	-9.79%	9.78%	19.54%
Std. dev. of PVC ($\$10^7$)	\$7.83148	\$8.69115	\$10.54365	\$11.50406
CV of PVC	0.02082	0.02060	0.02053	0.02057
Average LCT (year)	5.4177	5.9213	6.3948	6.2837
Change rate from base	-6.71%	1.96%	10.12%	8.20%
Std. dev. of LCT (year)	0.2186	0.2834	0.2736	0.2631
CV of LCT	0.04035	0.04785	0.04279	0.04188

By only changing all values of project construction cost (C_l, see Table 14) from the base case by the same percentage, four modified cases are composed, and the results obtained under these cases are compared in Table 24. Since the construction cost accounts for a very small fraction of PVC, the minimized expected PVC increases slightly with C_l. The optimized improvement sequence, however, is rather sensitive to C_l whose higher value discourages more projects from being implemented. The CV of PVC remains almost constant. With the increase in C_l, its effects of delaying project completion and disfavoring project implementation deal counter impacts on the average LCT.

Table 24 Optimized results with changes in the construction cost

Construction cost change	-20%	-10%	10%	20%
Optimized sequence	(7, 8, 9, 10, 1, 5)	(7, 8, 9, 10, 1, 5)	(7, 10, 8, 9, 1)	(7, 8, 10, 9)
Expected PVC ($\$10^9$)	\$4.64904	\$4.66341	\$4.69171	\$4.70457
Change rate from base	-0.61%	-0.30%	0.30%	0.58%
Std. dev. of PVC($\$10^7$)	\$9.61429	\$9.59851	\$9.65166	\$9.77814
CV of PVC	0.02068	0.02058	0.02057	0.02078
Average LCT (year)	5.3711	5.9422	6.3455	6.0859
Change rate from base	-7.51%	2.32%	9.27%	4.80%
Std. dev. of LCT (year)	0.2289	0.2568	0.3117	0.2658
CV of LCT	0.04261	0.04322	0.04913	0.04367

From the three tables above, it is noted that the CV of LCT appears to be more related to improvement sequences rather than to the average LCT – the same sequence with varying average LCTs across modified numerical cases has CVs of LCT at the same level, while different sequences have more dispersed CVs even when their average LCTs are closely distributed.

6.3.2 Sensitivity regarding the correlated uncertainties

6.3.2.1 With changed standard deviations

To explore the sensitivity of results to the correlated uncertainties, three modified numerical cases are generated. In each case, the standard deviations of one of the three correlated random variables, namely σ_g, σ_F, and σ_ψ, is doubled in the multivariate normal distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ as specified in Section 6.1.2. All other parameters are unchanged in each modified case. The optimization results are compared in Table 25. While the optimized sequence and its expected PVC are almost unchanged from the base case, the standard deviation and the CV of PVC are slightly more than doubled (due to congestion effects) when the standard deviation of the annual demand growth

rate (g) is doubled. The average LCT is most affected by the doubled standard deviation of the annual external budget supply (F), but it is only 1.20% higher than that in the base case. The doubled value of σ_F also causes the standard deviation and CV of LCT to double, while the doubled σ_ψ has slight impact on these two statistics. The reason is that in the proposed model the variation in F directly affects the starting and completion times of all projects in an improvement sequence. The impacts on completion times accumulate from the first to the last project in the sequence, making LCT the most affected completion time. However, the variation in ψ (the multiplier of required construction time) only affects the completion times of some projects in some scenarios. Such impacts do not accumulate over time and may counteract the impacts from the variation in F.

Table 25 Optimized results with doubled standard deviations of uncertain parameters

Changed std. dev. in $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	σ_g=0.005	σ_F=\$2.0×10^6	σ_ψ=0.2
Optimized sequence	(7, 8, 9, 10, 1)	(7, 8, 9, 10, 1)	(7, 10, 8, 9, 1)
Expected PVC (×\$10^9)	4.68231	4.67845	4.67840
Change rate from base case	0.100%	0.0179%	0.0168%
Std. dev. of PVC (×\$10^7)	19.54000	9.50248	9.64992
CV of PVC	0.04173	0.02031	0.02063
Average LCT (year)	5.8086	5.8769	5.8298
Change rate from base case	0.0214%	1.20%	0.387%
Std. dev. of LCT (year)	0.2743	0.5603	0.2632
CV of LCT	0.04723	0.09534	0.04515

6.3.2.2 With changed correlation coefficients

The three correlation coefficients (ρ_{gF}, $\rho_{g\psi}$, and $\rho_{F\psi}$) are changed from the base case to create five modified numerical cases. With the resulting changes in the correlation matrix $\boldsymbol{\Sigma}$, the vectors of uncertain parameters in the 50 scenarios differ. The

optimization results in each modified case are compared in Table 26. In the last case all three uncertain parameters are independent.

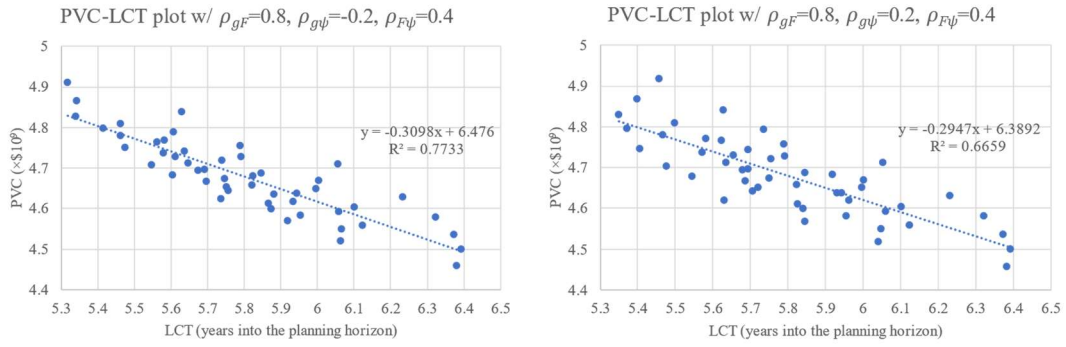
Table 26 Optimized results with modified correlation coefficients

Value of ρ_{gF}	0.8	0.4	0.8	0.9	0
Value of $\rho_{g\psi}$	-0.2	-0.2	0.2	0.6	0
Value of $\rho_{F\psi}$	0.4	0.2	0.4	0.7	0
Optimized sequence	(7, 8, 9, 10, 1)	(7, 8, 9, 10, 1)	(7, 8, 9, 10, 1)	(7, 8, 9, 10, 1)	(7, 8, 9, 10, 1)
Expected PVC (×$10^9)	4.67764	4.67762	4.67771	4.67787	4.67765
Change rate from base case	0.00056%	0.00020%	0.00208%	0.00555%	0.00082%
Std. dev. of PVC (×$10^7)	9.59373	9.71178	9.71348	9.78624	9.86433
CV of PVC	0.02051	0.02076	0.02077	0.02092	0.02109
Average LCT (year)	5.8055	5.8072	5.8069	5.8085	5.8065
Change rate from base case	-0.0323%	-0.0029%	-0.0077%	0.0199%	-0.0155%
Std. dev. of LCT (year)	0.2723	0.2654	0.2689	0.2615	0.2684
CV of LCT	0.04691	0.04570	0.04631	0.04502	0.04622

It appears that the differences in the optimized results are rather minor across different cases of parameter correlation. The optimized improvement sequences are the same as in the base case, with values of minimized expected PVC and average LCT being very close to those in the base case. The CVs of PVC and LCT are also at the same level as in the base case. However, when the values of PVC and LCT across 50 scenarios under each correlation case are plotted and compared in Figure 12, something interesting is observed. When all three uncertain parameters are independent, PVC and LCT can be treated as independent. With $\rho_{g\psi}$=-0.2 and larger coefficients of ρ_{gF} and $\rho_{F\psi}$, the correlation between PVC and LCT becomes stronger, as shown in the three sub-figures on the left of Figure 12. The coefficient ρ_{gF} has the main contribution to this correlation for the following reasons. Across the 50 scenarios, a higher annual growth rate of demand (g) significantly increase the cumulative travel time cost and

advances the starting and completion of projects by increasing the rate of internal budget supply. As pointed out in the above section, a higher annual external budget supply (F) has a stronger impact than a smaller multiplier of required construction time (ψ) on advancing the LCT. While the absolute values of $\rho_{g\psi}$ and $\rho_{F\psi}$ remain small (indicating weak g-ψ and F-ψ correlations), a larger and positive ρ_{gF} means that a higher g is more likely to be accompanied by a higher F, which means that a higher PVC is more likely to be accompanied by an earlier LCT.

When the sign of $\rho_{g\psi}$ is reversed, with the other two coefficients unchanged, the correlation between PVC and LCT is mildly weakened, as the two sub-figures at the top of Figure 12 show. With a positive $\rho_{g\psi}$ a higher g makes ψ more likely to be higher rather than lower, and a higher ψ has a contrary impact than a higher g on the LCT. Given a positive and large ρ_{gF}, positive and larger $\rho_{g\psi}$ and $\rho_{F\psi}$ weaken the PVC-LCT correlation. However, when comparing the second and the fourth sub-figure, it is observed that the increase in ρ_{gF} from 0.8 to 0.9 outweighs the increase in $\rho_{g\psi}$ and $\rho_{F\psi}$ with a strengthened PVC-LCT correlation.



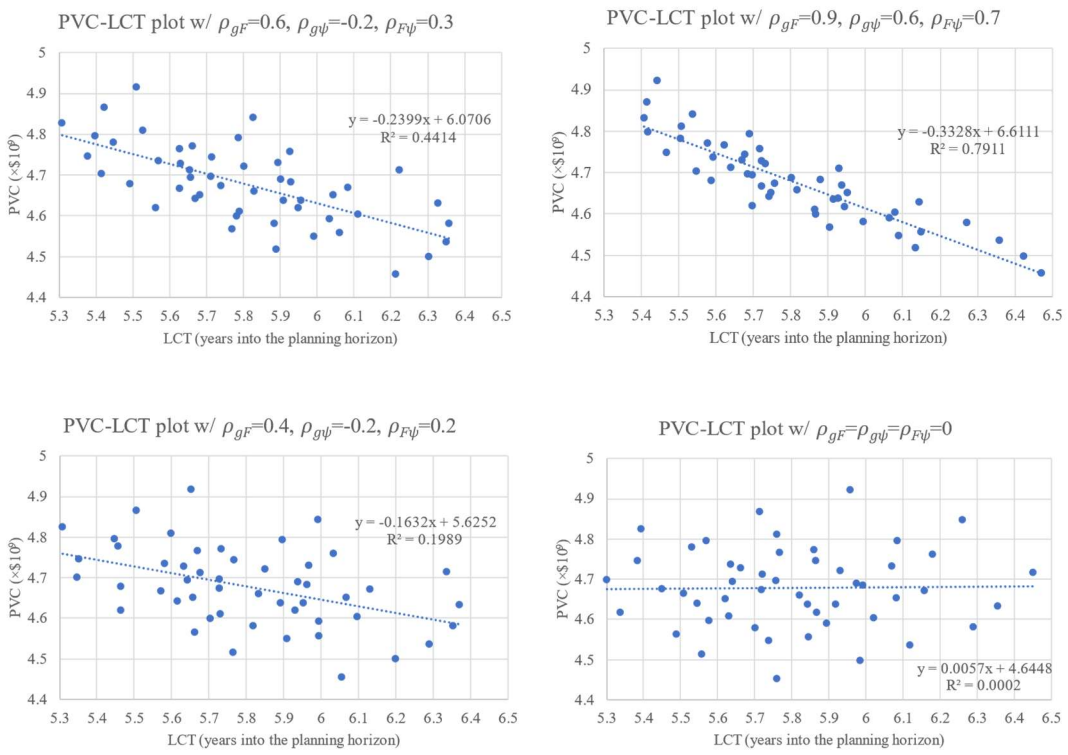


Figure 12 Comparing the correlation between PVC and LCT

6.4 Comparison to other improvement plans

6.4.1 Comparison to the optimized plan in the deterministic case

For comparison to the robustly optimized result that covers multiple scenarios, the selection, sequence, and schedule of improvement projects are optimized in numerical cases with deterministic values of uncertain variables. 7 cases shown in Table 27 are considered, among which 6 cases use the values of g, F, and ψ as in scenarios No. 1, 5, 15, 36, 46, 50 (listed in Table 15), and one case uses μ_g, μ_F, and μ_ψ specified in Section 6.1.2. All other parameters are the same as in the base case. The optimized results are compared in Table 27.

Table 27 Optimized results in each deterministic case of uncertain parameters

Based on	Scenario 1	Scenario 5	Scenario 15	μ_g, μ_F, μ_ψ	Scenario 36	Scenario 46	Scenario 50

$g(\%)$	1.918	2.165	2.362	2.500	2.638	2.835	3.082
$F(\$10^7)$	1.360	1.445	1.590	1.500	1.447	1.571	1.597
ψ	1.011	1.107	1.038	1.000	0.796	0.950	1.016
Opt. seq.	(7,8,9,10,1)	(7,8,9,10,1)	(7,8,9,10,1)	(7,8,9,10,1)	(7,10,8,9,1,5)	(7,8,9,10,1,5)	(7,8,9,10,1,5)
PVC ($\times\$10^9$)	4.45669	4.54909	4.61986	4.67556	4.72816	4.80825	4.91311
LCT (year)	6.21399	5.99278	5.56284	5.75458	6.49114	6.23188	6.20150

With a higher growth rate of demand, the savings in cumulative PV of travel time cost by completing an improvement project increase, and the implementation of more projects is favored. If the optimized sequence (7,8,9,10,1) in the base uncertain case is applied in the three deterministic cases based on scenarios No. 36, 46, and 50, the PVC will increase by $\$7.95\times10^5$(0.0168%), $\$5.53\times10^5$(0.0115%), and $\$8.71\times10^5$(0.0177%), respectively. If the optimized sequences in these three deterministic cases, namely (7,10,8,9,1,5) and (7,8,9,10,1,5), are applied in the base uncertain case, the resulting expected PVC increase by $\$6.77\times10^5$(0.0145%) and $\$6.00\times10^5$(0.0128%), respectively, compared to using the sequence (7,8,9,10,1).

6.4.2 Comparison to the greedy improvement plan

The greedy algorithm for optimizing the selection and sequencing of improvement projects is performed through the following process. Before the iteration starts, the expected PVC without any improvement is computed as the "base value". In the first iteration, each project is chosen out of all ten candidates, and the expected PVC (excluding construction cost, in all iterations) is computed for each of all ten improvement plans where only the chosen project is implemented. The ratio of incremental benefit ("base value" minus the newly computed expected PVC) to

incremental cost (average PV of construction cost of the chosen project) is computed for each chosen project. If the highest ratio is above 1, the project with that highest ratio becomes the first project to be implemented in the sequence. Its corresponding expected PVC is recorded as the new "base value". In the next iteration, each project from the remaining nine candidates is evaluated as the second project to be implemented, and the expected PVC along with the ratio specified above is computed for each of the resulting sequences. If the highest ratio is above 1, the corresponding project is determined as the second project in the sequence, with the "base value" updated to the corresponding expected PVC before proceeding to the third iteration. Such iterations continue with projects added to the improvement sequence one by one, until all projects are included in the improvement sequence or the search is terminated due to the highest ratio decreasing below 1.

With all parameters unchanged from the base case, the changes in benefit-cost ratios with respect to each additional project to be implemented are shown in Table 28. The bolded numbers are the highest ratios in each iteration, indicating the project to be added to the optimized improvement sequence at each step.

Table 28 Benefit-cost ratios in each iteration of the greedy algorithm

Added Proj. #	1	2	3	4	5	6	7	8	9	10
Iter. 1	2.359	1.665	1.224	1.615	2.382	0.245	**7.232**	2.984	2.754	2.520
Iter. 2	2.164	1.298	0.807	1.243	1.882	0.104		**2.785**	1.851	2.533
Iter. 3	1.850	1.188	0.500	1.076	2.049	0.069			2.469	**2.503**
Iter. 4	1.535	0.823	0.354	0.689	1.480	-0.102			**2.936**	
Iter. 5	**1.262**	0.627	0.035	0.691	1.003	-0.276				
Iter. 6		0.571	0.015	0.609	0.916	-0.330				

The improvement sequence optimized by the greedy algorithm is represented by (7,8,10,9,1), which differs slightly from the GA-optimized sequence of (7,8,9,10,1). Its expected PVC is $\$4.67771\times10^9$, higher than the GA-minimized PVC by a margin of only $\$9.70\times10^4$ (0.00207%), which indicates that in this case myopic search is almost as effective as a more sophisticated search. This may be attributable to the relatively small number of candidate projects, the non-overlapping affected links of these projects, and the relatively small number of OD pairs with non-zero demand (compared to the number of all possible OD pairs) with a low congestion level in the road network.

Chapter 7: Conclusion and possible improvements

The work in this dissertation focuses on joint optimizations involving selection, sequencing, and scheduling of interrelated projects in transportation networks with uncertainties. Two problems are discussed: one for rail freight networks, and the other for road networks.

For rail freight networks, a tri-level model that integrates optimizations of short-term post-disruption restoration sequencing and scheduling and long-term improvement project selection and scheduling is developed. An approximation method is proposed for considering long-term demand growth in the computation of cumulative expected excess. A genetic algorithm (GA) is applied for optimization searches with larger numbers of damaged components or improvement projects. The optimized sequences are translated into schedules under the rules based on binding resource and budget constraints. The proposed model is demonstrated with short-term and long-term numerical cases consisting of a small demand-loaded network, a set of disruption scenarios, and a list of improvement projects. In the short-term problem, with different numbers of available work teams, the globally optimal restoration sequences (as shown through exhaustive enumeration) are obtained by the GA in a few seconds. The restoration itineraries and schedules of work teams along with corresponding changes in hourly cost are shown in figures. In the long-term problem, optimal selections and sequences of improvement projects are obtained through analysis of sensitivity to the number of work teams, planning horizon duration, and project construction cost. Availability of more work teams greatly reduces cumulative expected excess by shortening the duration of post-disruption restoration, and therefore discourages

implementation of long-term improvement projects. With a longer planning horizon, lower construction costs, and a higher interest rate, more long-term improvements are favored. Evaluations of cumulative expected excess with probabilistic growth rates of demand are provided. The model is also tested on a larger-scale real-world-based network, for which the effectiveness of GA in optimizing the long-term improvement plan is verified through a statistical test.

For road networks, a bi-level model is designed for jointly assigning traffic to user equilibrium (UE) at the lower level and optimizing the selection and sequencing of network improvement projects at the upper level. To consider correlated uncertainties at the upper level, multiple scenarios are generated with samples from a multivariate distribution of uncertain parameters. Under each scenario, an improvement sequence is mapped to a unique schedule based on binding budget and work duration constraints. The planning horizon is divided into short sub-periods to approximate the effects of demand growth and cost discounting. The lower-level model computes the average hourly travel time cost in each sub-period under the time-varying network configuration, and the upper-level model minimizes the expected present value of cost (PVC, cumulative travel time cost plus project construction cost) over the planning horizon. A set of methods is proposed for determining budget-ready times of projects with internal budget supply enabled. The model also allows the use of two modes – bus and car – in the network, and the iteration of mode shares is integrated with the lower-level F-W traffic assignment. In the numerical case, the three correlated uncertain parameters are demand growth rate, external budget supply, and the multiplier of required construction time of projects. 50 scenarios are generated with an underlying

Hammersley sequence in the process of vector sampling. With the help of parallelized computation, the GA-optimized selection and sequencing of improvement projects is obtained within a day. Under the optimized sequence, the PVCs and the last completion times (LCTs) in each scenario are listed. The optimized results, including the optimized sequence and the key statistics of PVC and LCT across all scenarios, are obtained and compared through analysis of sensitivity to the use of bus, demand level, users' value of travel time, construction cost, and the standard deviations and correlation coefficients of uncertain parameters. When introducing buses to the network, two operation modes are considered: using dedicated bus lanes and sharing roads with cars. The iteration of bus shares shows that the bus operation is more desirable with dedicated lanes, and that the potential demand level of bus passengers is vital to the desirability. The implementation of more projects is favored by a higher demand level, a higher value of travel time, and lower construction costs. Across the scenarios, the demand growth rate and the external budget supply have direct impacts on PVC and LCT, respectively. Therefore, the correlation between PVC and LCT varies with different correlation matrices of uncertain parameters. The optimized results in the base case are compared to those obtained in deterministic cases or with a greedy algorithm, from which a higher demand growth rate is also found to favor project implementation.

Possible future extensions of this work include the following:

In Problem 1:

1) Multiple types of freight may be considered, with varied unit shipment costs, shipment priorities, and values of time. In this context capacity constraints are applied to total amount of freight shipped through each component.

2) More details can be applied to representations of disruption scenarios. Damage patterns and probabilities of occurrence can be specified for each type of disruption at each network component. A link may be divided into segments with varied degrees of damage and, correspondingly, varied unit restoration time and cost.
3) More flexible options can be considered for short-term restoration, with its cost included in the excess. Faster restoration with higher unit costs, partial restoration, and pre-positioning of constrained resources (work teams, equipment, and materials) can be considered in heuristic optimization.
4) Temporary storage and delay of freight at stations (nodes) or trains may be considered, with their costs included in the excess. This may involve conservation of locomotives and cars.
5) Starting and completion times of long-term improvement projects may become optimizable under constraints from available budget and future demand levels. Borrowing and lending among projects may be allowed in using the budget.

In Problem 2:

1) To make the model more applicable to urban road networks, intersection and signal delays may be considered in the F-W traffic assignment.
2) If computation resources permit, traffic simulation methods such as cell transmission and link transmission may be applied to provide more accurate results of total travel time.

3) The buses may be allowed to use dedicated lanes in some routes and share roads with cars in others. More cost items and coefficients may be included in the impedance functions.
4) More general forms of budget accumulation, demand growth, and multivariate distributions, considering more uncertainties, may be included.
5) The effects of demand elasticities on equilibrium flows should be included in future studies.
6) In the numerical case with correlated uncertainties, the variation ranges of uncertain parameters are considerably smaller than some which arise in the real-world cases. In future work, larger coefficients of variation (CVs) of uncertain parameters as well as more dispersed distributions should be explored in numerical cases so that the advantages of optimizing solutions under uncertainties can be more clearly shown.
7) A neural network might be trained to estimate average hourly travel time cost from the input parameters, so that the total computation time may be reduced significantly.

References

Akbari, V., Sadati, M. E. H., & Kian, R., 2021a. A decomposition-based heuristic for a multicrew coordinated road restoration problem. *Transportation Research Part D: Transport and Environment*, 95, 102854.

Akbari, V., Shiri, D., & Salman, F. S., 2021b. An online optimization approach to post-disaster road restoration. *Transportation Research Part B: Methodological*, 150, 1-25.

Aksu, D. T., & Ozdamar, L., 2014. A mathematical model for post-disaster road restoration: Enabling accessibility and evacuation. *Transportation Research Part E: Logistics and Transportation Review*, 61, 56-67.

Almoghathawi, Y., Barker, K., & Albert, L.A., 2019. Resilience-driven restoration model for interdependent infrastructure networks. *Reliability Engineering & System Safety*, 185, 12-23.

Ayyub, B.M., 2014. Systems resilience for multihazard environments: Definition, metrics, and valuation for decision making. *Risk analysis*, 34(2), 340-355.

Ayyub, B.M., 2015. Practical resilience metrics for planning, design, and decision making. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A*: Civil Engineering, 1(3), p.04015008.

Bagloee, S. A., & Asadi, M., 2015. Prioritizing road extension projects with interdependent benefits under time constraint. *Transportation Research Part A: Policy and Practice*, 75, 196-216.

Bouleimen, K., & Lecocq, H., 2003. A new efficient simulated annealing algorithm for the resource-constrained project scheduling problem and its multiple mode version. *European journal of operational research*, 149(2), 268-281.

Bureau of Public Roads, 1964. *Traffic Assignment Manual*. US Department of Commerce, Urban Planning Division, Washington, D.C.

Cacchiani, V., Huisman, D., Kidd, M., Kroon, L., Toth, P., Veelenturf, L., & Wagenaar, J., 2014. An overview of recovery models and algorithms for real-time railway rescheduling. *Transportation Research Part B: Methodological*, 63, 15-37.

Chen, H., Cullinane, K., & Liu, N., 2017. Developing a model for measuring the resilience of a port-hinterland container transportation network. *Transportation Research Part E: Logistics and Transportation Review*, 97, 282-301.

Chen, L., & Miller-Hooks, E., 2012. Resilience: an indicator of recovery capability in intermodal freight transport. *Transportation Science*, 46(1), 109-123.

Dao, C.D., Hartmann, A., Lamper, A., & Herbert, P., 2019. Scheduling infrastructure renewal for railway networks. *Journal of Infrastructure Systems*, 25(4), p.04019027.

Erlenkotter, D., 1973. Sequencing of interdependent hydroelectric projects. *Water Resources Research*, 9(1), 21-27.

Faturechi, R., & Miller-Hooks, E., 2014. Travel time resilience of roadway networks under disaster. *Transportation Research Part B: Methodological*, 70, 47-64.

Frank, M., & Wolfe, P., 1956. An algorithm for quadratic programming. *Naval research logistics quarterly*, 3(1-2), 95-110.

Goldberg, D.E., & Lingle, R., 1985, July. Alleles, loci, and the traveling salesman problem. In *Proceedings of an international conference on genetic algorithms and their applications*, 154, 154-159. Hillsdale, NJ: Lawrence Erlbaum.

Gong, L., & Fan, W., 2016. Optimizing scheduling of long-term highway work zone projects. *International journal of transportation science and technology*, 5(1), 17-27.

Gu, Y., Fu, X., Liu, Z., Xu, X., & Chen, A., 2020. Performance of transportation network under perturbations: Reliability, vulnerability, and resilience. *Transportation Research Part E: Logistics and Transportation Review*, 133, 101809.

Hosseininasab, S. M., Shetab-Boushehri, S. N., Hejazi, S. R., & Karimi, H., 2018. A multi-objective integrated model for selecting, scheduling, and budgeting road construction projects. *European Journal of Operational Research*, 271(1), 262-277.

Hu, Y. C., & Schonfeld, P., 1984. Simulation and optimization of regional road network. *Journal of transportation engineering*, 110(4), 431-443.

Jin, J. G., Tang, L. C., Sun, L., & Lee, D. H., 2014. Enhancing metro network resilience via localized integration with bus services. *Transportation Research Part E: Logistics and Transportation Review*, 63, 17-30.

Jong, J. C., & Schonfeld, P., 2001. Genetic algorithm for selecting and scheduling interdependent projects. *Journal of waterway, port, coastal, and ocean engineering*, 127(1), 45-52.

Jovanovic, U., Shayanfar, E., & Schonfeld, P., 2018. Selecting and Scheduling Link and Intersection Improvements in Urban Networks. *Transportation Research Record*, 2672(51), 1-11.

Kalagnanam, J. R., & Diwekar, U. M., 1997. An efficient sampling technique for off-line quality control. *Technometrics*, 39(3), 308-319.

Karlaftis, M.G., Kepaptsoglou, K.L., & Lambropoulos, S., 2007. Fund allocation for transportation network recovery following natural disasters. *Journal of Urban Planning and Development*, 133(1), 82-89.

Kasaei, M., & Salman, F. S., 2016. Arc routing problems to restore connectivity of a road network. *Transportation Research Part E: Logistics and Transportation Review*, 95, 177-206.

Kepaptsoglou, K.L., Konstantinidou, M.A., Karlaftis, M.G., & Stathopoulos, A., 2014. Planning postdisaster operations in a highway network: Network design model with interdependencies. *Transportation Research Record*, 2459(1), 1-10.

Kumar, A., & Peeta, S., 2014. Slope-based path shift propensity algorithm for the static traffic assignment problem. *International Journal for Traffic and Transport Engineering*, 4(3), 297-319.

Kumar, A., & Mishra, S., 2017. A simplified framework for sequencing of transportation projects considering user costs and benefits. *Transportmetrica A: Transport Science*, 14(4), 346-371.

LeBlanc, L. J., Morlok, E. K., & Pierskalla, W. P., 1975. An efficient approach to solving the road network equilibrium traffic assignment problem. *Transportation research*, 9(5), 309-318.

Li, Z., Roshandeh, A. M., Zhou, B., & Lee, S. H., 2013. Optimal decision making of interdependent tollway capital investments incorporating risk and uncertainty. *Journal of transportation engineering*, 139(7), 686-696.

Liu, J., Schonfeld, P., Peng, Q., & Yin, Y., 2020a. Measures of travel reliability on an urban rail transit network. *Journal of Transportation Engineering, Part A: Systems*, 146(6), p.04020037.

Liu, J., Schonfeld, P., Yin, Y., & Peng, Q., 2020b. Effects of Link Capacity Reductions on the Reliability of an Urban Rail Transit Network. *Journal of Advanced Transportation*, 2020.

Liu, K., Zhai, C., & Dong, Y., 2021. Optimal restoration schedules of transportation network considering resilience. *Structure and Infrastructure Engineering*, 17(8), 1141-1154.

Mattsson, L.G., & Jenelius, E., 2015. Vulnerability and resilience of transport systems–A discussion of recent research. *Transportation Research Part A: Policy and Practice*, 81, 16-34.

Miandoabchi, E., Daneshzand, F., Zanjirani Farahani, R., & Szeto, W. Y., 2015. Time-dependent discrete road network design with both tactical and strategic decisions. *Journal of the Operational Research Society*, 66(6), 894-913.

Mika, M., Waligóra, G., & Węglarz, J., 2005. Simulated annealing and tabu search for multi-mode resource-constrained project scheduling with positive discounted cash flows and different payment models. *European Journal of Operational Research*, 164(3), 639-668.

Miller-Hooks, E., Zhang, X., & Faturechi, R., 2012. Measuring and maximizing resilience of freight transportation networks. *Computers & Operations Research*, 39(7), 1633-1643.

Miralinaghi, M., Seilabi, S. E., Chen, S., Hsu, Y. T., & Labi, S., 2020a. Optimizing the selection and scheduling of multi-class projects using a Stackelberg framework. *European Journal of Operational Research*, 286(2), 508-522.

Miralinaghi, M., Woldemariam, W., Abraham, D. M., Chen, S., Labi, S., & Chen, Z., 2020b. Network-level scheduling of road construction projects considering user and business impacts. *Computer-Aided Civil and Infrastructure Engineering*, 35(7), 650-667.

Mohammadi, R., He, Q., & Karwan, M., 2020. Data-driven robust strategies for joint optimization of rail renewal and maintenance planning. *Omega*, 103, p.102379.

Nemhauser, G. L., & Ullmann, Z., 1969. Discrete dynamic programming and capital allocation. *Management Science*, 15(9), 494-505.

Ng, M., & Schonfeld, P., 2023. Sequencing interdependent transportation network improvement projects: exact solution method via network flow reformulation, *Transportation Research Part D: Transport and Environment,* v115 103565.

Peng, Y.T., Li, Z.C., & Schonfeld, P., 2019. Development of rail transit network over multiple time periods. *Transportation Research Part A: Policy and Practice*, 121, 235-250.

Saadat, Y., Ayyub, B.M., Zhang, Y., Zhang, D., & Huang, H., 2019. Resilience of metrorail networks: Quantification with Washington, DC as a case study. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, 5(4).

Sharkey, T.C., Cavdaroglu, B., Nguyen, H., Holman, J., Mitchell, J.E., & Wallace, W.A., 2015. Interdependent network restoration: On the value of information-sharing. *European Journal of Operational Research*, 244(1), 309-321.

Shayanfar, E., & Schonfeld, P., 2018. Selecting and scheduling interrelated projects: application in urban road network investment. *International Journal of Logistics Systems and Management*, 29(4), 436-454.

Shayanfar, E., & Schonfeld, P., 2019. Selecting and scheduling interrelated road projects with uncertain demand. *Transportmetrica A: Transport Science*, 15(2), 1712-1733.

Sun, J., & Zhang, Z., 2020. A post-disaster resource allocation framework for improving resilience of interdependent infrastructure networks. *Transportation Research Part D: Transport and Environment*, 85, p.102455.

Tao, X., & Schonfeld, P., 2006. Selection and scheduling of interdependent transportation projects with island models. *Transportation research record*, 1981(1), 133-141.

Tao, X., & Schonfeld, P., 2007. Island models for stochastic problem of transportation project selection and scheduling. *Transportation research record*, 2039(1), 16-23.

Tofighian, A. A., & Naderi, B., 2015. Modeling and solving the project selection and scheduling. *Computers & Industrial Engineering*, 83, 30-38.

Vugrin, E. D., Turnquist, M. A., & Brown, N. J., 2014. Optimal recovery sequencing for enhanced resilience and service restoration in transportation networks. *Int. J. Crit. Infrastructures*, 10(3/4), 218-246.

Wang, S. L., & Schonfeld, P., 2005. Scheduling interdependent waterway projects through simulation and genetic optimization. *Journal of waterway, port, coastal, and ocean engineering*, 131(3), 89-97.

Wang, S. L., & Schonfeld, P., 2008. Scheduling of waterway projects with complex interrelations. *Transportation Research Record*, 2062(1), 59-65.

Wardrop, J. G., 1952. Road paper. some theoretical aspects of road traffic research. *Proceedings of the institution of civil engineers*, 1(3), 325-362.

Wei, C. H., & Schonfeld, P., 1994. Multiperiod network improvement model. *Transportation Research Record*, 110-110.

Whitley, L.D., 1989. The GENITOR algorithm and selection pressure: why rank-based allocation of reproductive trials is best. *Icga*, 89, 116-123.

Woodburn, A., 2019. Rail network resilience and operational responsiveness during unplanned disruption: A rail freight case study. *Journal of Transport Geography*, 77, 59-69.

Wu, F., Schonfeld, P., & Kim, M., 2021. Optimized restoration schedule for disrupted railroad network. *Journal of Transportation Engineering, Part A: Systems*, 147(9), p.04021053.

Zhang, D.M., Du, F., Huang, H., Zhang, F., Ayyub, B.M., & Beer, M., 2018. Resiliency assessment of urban rail transit networks: Shanghai metro as an example. *Safety Science*, 106, 230-243.

Zhang, L., Wen, Y., & Jin, M., 2009. The framework for calculating the measure of resilience for intermodal transportation systems.

Zhang, L., Wen, Y., Huang, Z., & Jin, M., 2010. Framework of calculating the measures of resilience (MOR) for intermodal transportation systems. (No. FHWA/MS-DOT-RD-10-220). Mississippi. Dept. of Transportation.