Vulnerability Assessment and Mitigation for the Chinese Railway System under Floods

Liu Hong a,b,c, Min Ouyang a,c*, Srinivas Peeta b,d, Xiaozheng He b, Yongze Yan a

a School of Automation, Huazhong University of Science and Technology, Wuhan 430074, China
b NEXTRANS Center, Purdue University, W. Lafayette, IN 47906, USA
c Key Lab for Image Processing and Intelligent Control, Huazhong University of Science and Technology, Wuhan 430074, China
d School of Civil Engineering, Purdue University, W. Lafayette, IN 47906, USA

Abstract

The economy of China and the travel needs of its citizens depend significantly on the continuous and reliable services provided by its railway system. However, this system is subject to frequent natural hazards, such as floods, earthquakes, and debris flow. A mechanism to assess the railway system vulnerability under these hazards and the design of effective vulnerability mitigation strategies are essential to the reliable functioning of the railway system. This paper proposes a comprehensive methodology to quantitatively assess the railway system vulnerability under floods using historical data and GIS technology. The proposed methodology includes a network representation of the railway system, the generation of flood event scenarios, a method to estimate railway link vulnerability, and a quantitative vulnerability value computation approach. The railway system vulnerability is evaluated in terms of its service disruption related to the number of interrupted trains and the durations of interruption. A maintenance strategy to mitigate vulnerability is proposed that simultaneously considers link vulnerability and number of trains using it. Numerical experiments show that the flood-induced vulnerability of the proposed representation of the Chinese railway system reaches its maximum monthly value in July, and the proposed vulnerability mitigation strategy is more effective compared to other strategies. The methodology can be extended to assess the railway system vulnerability under other natural hazards. It is further helpful for identifying an effective way to distribute limited maintenance resources to a transportation network facing frequent natural hazards.

Key words: Chinese railway system; Network; Flood hazard; Vulnerability assessment; Vulnerability mitigation;

* Corresponding author. Tel.: +86-27-87559490; fax: +86-27-87543130.

E-mail addresses: liu.hong@hust.edu.cn (L. Hong), min.ouyang@hust.edu.cn (M. Ouyang), peeta@purdue.edu (S. Peeta), ychyzz@163.com (Y. Yan), seanhe@purdue.edu (X. He).
1. Introduction

Railway systems have played a key role in the economic and social development in many countries since the nineteenth century. In 2012, about 1.89 billion passengers and 3.89 billion metric tons of cargo were transported by the Chinese railway system (CRS) [1]. Here, CRS includes both the regular and high-speed railway networks, whose links typically run in parallel. Disruptions in a railway system can have severe consequences, such as direct damage and indirect loss [2]. Floods represent one of the most important natural hazards, and cause at least one-third of the total losses due to all natural hazards in the world [3]. China is a country prone to flood hazards. Two-thirds of the Chinese land area faces the threat of floods, which can disrupt the CRS and lead to enormous economic losses. From 2000 to 2010, the average spending on flood-related maintenance in the CRS was about $15 million per year, and the indirect loss caused by railway disruption was about $37.5 million per year [1]. Hence, an assessment of the vulnerability of the CRS under floods and the design of effective vulnerability mitigation strategies are critical for the reliable functioning of the CRS and the protection of its infrastructure.

Studies from different domains define system vulnerability differently [4-8]. As a synthesis of the available literature, this paper considers that the vulnerability is associated with a disruptive event and quantified by the performance drop of the system under the event. For the vulnerability analysis of infrastructure systems, many approaches exist in the literature to address this issue [9], such as agent based approaches [10-15], system dynamics based approaches [16-20], network based approaches [21-25], and others. The agent based approaches use agents to represent components in an infrastructure (such as electric transformers or generators) or important players (such as government or weather) related to system operation. A set of rules is used in agent-based approaches to describe agents’ behaviors and their interactions with the environment and capture system performance response under disruptive events. Discrete-event simulations are used in agent based approaches to provide scenario-based vulnerability analysis and the effectiveness assessment of different vulnerability mitigation strategies. The system dynamics approaches use feedback loops, stocks, and flows to model the dynamic and evolutionary behavior of infrastructure systems under disruptive scenarios to analyze the effects of policy and technique factors on system evolution in the long term and provide investment recommendations to mitigate vulnerability. The network based approaches model each involved system as a network, which enables capturing the topological features of infrastructure systems and flow characteristics, identifying the critical system components, and providing suggestions on mitigation strategies at
detailed topological levels. As the Chinese railway system is distributed over a large-scale area and the flood hazards are affected by the geographical location, modeling the flood impact on the railway system requires describing the railway system at a topological level. Hence, this paper uses a network-based approach to model flood-induced railway vulnerability.

To assess the railway system vulnerability at the topological level, some studies analyze the topological characteristics of railway systems of various countries (India [26], Spain [27] and China [28]) using complex network theory [29, 30]. They describe the railway system through a graph, with nodes representing railway stations and links representing the relationships between stations. Using this method, Wang et al. [31] discuss the topological properties of the CRS using two models and illustrate that the CRS exhibits properties such as hierarchical structure and small-world behavior. Derrible and Kennedy [32] analyze the complexity and robustness of 33 metro systems in the world based on network science methodologies, and conclude that robustness is related to the number of cycles in the network. They also provide recommendations to increase the system robustness at a macro level. Zhang et al. [33, 34] discuss the vulnerability of the China high speed railway network and the Shanghai subway network. They use transport topological efficiency loss and connectivity to describe the vulnerability of a railway network under random and intentional attacks. These studies are usually aimed at random failures or intentional attacks, but there exists another type of disruptive events, called natural hazards, which disrupt the railway system in different ways. These hazards can attack several infrastructure components simultaneously, and each component may have specific failure probabilities corresponding to the type of natural hazard. Chang and Nojima [35] introduce a method to analyze the post-disaster performance of the railway system under earthquake scenarios, which is evaluated in terms of the total length of functional railway line and the total distance-based accessibility. Moran et al. [36] describe a documentation procedure to characterize the structural damage of railway infrastructure components and associated operational effects under the impact of floods.

While there is little work on system level vulnerability analysis of railway networks under natural hazards, there exist many network-based system-level vulnerability studies on other types of transportation and infrastructure systems under natural hazards. Nagurney et al. [37] propose environmental link importance indicators to describe the environmental impacts of the degradation in road network link capacities. Peeta et al. [38] assess the vulnerability of highway networks subject to random failures under earthquake, and propose a method to make pre-disaster investment decisions for strengthening the highway network. Jenelius and Mattson [39] present a
method to analyze the vulnerability of road networks under area-covering disruptions. In it, the road network is covered using a grid of uniformly shaped and sized cells, where each cell represents the spatial extent of a disrupting event. Sohn [40] uses an accessibility perspective to study the vulnerability of a highway network under flood damage by evaluating the significance of its links. More comprehensive vulnerability studies of various infrastructure systems under natural hazards include the seismic vulnerability analysis of power grid and water pipeline system in Shelby County, TN, USA [41-45], European gas and electricity systems [46], the hurricane vulnerability analysis of power grid and gas system in Harris County, TX, USA [47-50], and the lightning vulnerability of the IEEE 118 power grid [51].

The aforementioned vulnerability studies are aimed at different systems and different types of hazards, but illustrate a common modeling framework. This framework includes the following steps: (1) modeling the hazards of concern to generate a hazard scenario; (2) estimating component failure probabilities under the hazard scenario; (3) comparing each component failure probability with a uniformly distributed random number to produce a damage event which describes the damage state of each component; and (4) modeling and analyzing system performance response under the initial component damage or the specific event, and computing the system performance drop under the event, which is labeled the vulnerability under the event. The procedure is repeated under different events generated using the random number, and the average computed vulnerability value across the events is regarded as the vulnerability under that specific hazard. This paper applies this framework to propose an approach to quantitatively assess the vulnerability of a railway system under flood hazards by using historical flood data and geographic information systems (GIS) technology. This method consists of four parts that are illustrated using the CRS: (i) a network representation of the CRS is provided and some of its topological properties are discussed, (ii) flood event scenarios are generated through Monte Carlo simulation using historical flood event data for the past 30 years in China, (iii) the railway link vulnerabilities are estimated based on flood-induced railway disruption event data for the past 30 years, and (iv) the concept of railway service disruption is introduced and used to quantify the railway system vulnerability.

Also, note that for the railway vulnerability quantification, in the literature, many studies use purely railway topological models and metrics, but recent studies show flow is a key factor influencing system vulnerability [52]. Ouyang et al. [53] compare two complex network based models, including a purely topological model which does not consider train flow and a shortest path model which considers train flow running along the shortest paths, with a real train flow
model. Results in [53] show that the complex network based models can approximate the real flow model well to produce railway vulnerability under single component failures, but not well under multiple component failures. Natural hazards, such as floods and earthquakes, almost always affect large areas and many components of the system simultaneously. Hence, this paper uses a real train flow model to assess the flood-induced railway vulnerability.

The remainder of the paper is organized as follows. Section 2 introduces the vulnerability assessment methodology for railway systems under floods, including a network representation of the railway system, the generation of flood event scenarios, a method to estimate railway link vulnerability, and a quantitative system-level vulnerability value computation approach. Section 3 applies the proposed methodology to analyze the vulnerability of the CRS. It also proposes a maintenance strategy for the CRS vulnerability mitigation. Numerical examples are discussed to compare the effectiveness of various vulnerability mitigation strategies with the proposed one. Section 4 provides some concluding comments and potential directions for future research.

2. Vulnerability assessment of a railway system under floods

According to the literature review introduced in Section 1, the proposed approach includes a network representation of the railway system, the generation of flood event scenarios, a method to estimate railway link vulnerability, and a quantitative system-level vulnerability value computation model.

2.1 Network representation of railway system

The CRS consists of 2940 stations [54]. Among them, only a subset of stations is selected to construct a representative railway network for this study. They include stations located in or near cities with large population and GDP, and those that represent the start or end points of a train. Since the vulnerability assessment is determined using which trains use which stations/links (based on the data availability), the smaller stations are not explicitly included as nodes unless they represent the start or end points of trains. Thereby, the links between the stations included in the proposed network representation as nodes incorporate implicitly the corresponding smaller stations (along that railway route) that are not represented as nodes in the sense that the number of trains passing through a link is an accurate representation of the CRS train routes. Multiple stations located in the same city, such as in Beijing, Shanghai, and Wuhan, are combined as one for simplification. Based on these criteria, the proposed CRS network representation consists of 399 stations and 500 railway links as shown in Fig. 1 [54]. This CRS representation is denoted as a network \( G = (S, L) \), where \( S \) denotes the set of railway stations, and \( L \) denotes the set of railway
links connecting these stations. Also, this network representation of the CRS will be labeled as CRSN.

![Proposed network representation of the Chinese railway system.](image)

The topological properties of the CRSN are shown in Table 1, including the average degree (the average number of links to which a station is connected), the diameter (the maximum shortest path length between any pair of stations) and the mean distance (the average shortest path length between any pair of stations), where the shortest path length means the least number of links between a selected pair of stations. The topological properties of two other railway networks are listed in the table for comparison purpose. It illustrates that the three railway networks have similar connectivity level, with the average degree ranging from 2.1 to 2.5. However, the diameter and the mean distance of the CRSN are much smaller than those of the Swiss and European railway networks [55]. This is because the proposed network representation includes mostly the important stations in the CRS as discussed earlier.

**Table 1** Topological properties of the CRSN and other railway systems.

<table>
<thead>
<tr>
<th>Railway Networks</th>
<th>Number of Stations</th>
<th>Number of Links</th>
<th>Average Degree</th>
<th>Diameter</th>
<th>Mean Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>China (CRSN)</td>
<td>399</td>
<td>500</td>
<td>2.5</td>
<td>39</td>
<td>15.0</td>
</tr>
<tr>
<td>Swiss</td>
<td>1613</td>
<td>1680</td>
<td>2.1</td>
<td>136</td>
<td>46.6</td>
</tr>
<tr>
<td>European</td>
<td>4853</td>
<td>5765</td>
<td>2.4</td>
<td>184</td>
<td>50.9</td>
</tr>
</tbody>
</table>

As of 2010, the CRS had 4196 passenger trains running upon the system during a typical weekday [54]. Denote \( T = \{t\} \) as the set of trains. Then, the routes of all trains can be described
by a 4196 by 500 incidence matrix $R = \{r_{lt}\}$. If train $t$ passes through the railway link $l$, $r_{lt}=1$, otherwise, $r_{lt}=0$, where $t \in T, l \in L$.

2.2 Flood scenario generation

An analysis of the impact of floods on the railway system entails the representation of the associated flood patterns. There are many models for flood inundation prediction [56, 57], some of which have been used for many years and are well-accepted decision support tools. Liu and Matthies [58] model the flood in Toce river valley in Italy based on shallow water equations. Cummings et al. [59] use the Hazus-MH MR4 flood model to analyze potential flood damage in Minnewaukan, North Dakota in USA. However, most of these studies address regional area flood or specific river flood, and lack models for large area floods due to insufficient data [60]. The flood-threatened land of China is more than 6,400,000 KM$^2$. The flood type varies from region to region, such as river floods in flatlands, melting snow floods in high altitudes, and mountain floods. Different types of floods affect the railway system in different ways. Additionally, the occurrence of floods is affected by many factors such as rainfall per hour, geological conditions, and terrain situations. Currently, only the data of each historical flood event occurrence time and location is available. Hence, this paper uses a simple probabilistic model to generate flood scenarios. If detailed geographical and hydrological information is available, a more comprehensive flood scenario generation model can be employed.

To generate flood scenarios, flood event data in China from 1981 to 2010 was collected, including the occurrence time and the location of each flood event. These flood event data are grouped by their occurrence months and location (province). Then, the number of flood events for each province in each month is determined. This data set is illustrated in Fig. 2 which indicates
that, over the past 30 years, more than 90% of the flood events in China occurred in four months: June through September.

Fig. 3 indicates the total number of flood events in each province from June through September. The three provinces with the largest frequency of flood events in June are Hunan, Jiangxi and Fujian, and in July they are Hubei, Hunan and Guangxi. Fig. 3 illustrates that the flood event occurrence time and frequency vary with location. In most of the middle provinces, flood events occur frequently in June, July and August, as this period represents the rainy season in this area of China. By contrast, most coastal provinces have high flood event frequencies in August and September, as most typhoons occur in these two months. Based on the available monthly flood occurrence data, this study assumes that each day in a month has identical flood occurrence probability. The daily flood occurrence probability of a province in a particular month is computed as the ratio of the total number of historical flood events (over the 30 years) occurring in that province in that month to the product of the number of years (30, in this study) and the number of days in that month.

![Fig. 3. The flood event occurrence times in each province over the four months.](image)

A flood event scenario is generated using Monte Carlo simulation based on the daily flood occurrence probabilities of each province in a particular month. To generate a flood event scenario for a province, a uniformly distributed random number within [0, 1) is produced and compared with the daily flood occurrence probability of the province. If the random number is larger than or equal to that daily flood occurrence probability, the flood is assumed not to occur in that province; otherwise, the flood is assumed to occur. There are 31 provinces covered by
connected railway lines in China. The flood occurrences in the different provinces are assumed to be independent of each other. Then, a flood event scenario for China on a particular day is obtained after 31 such computations. Thereby, flood events may occur simultaneously in different provinces.

2.3 Estimation of the vulnerability of the railway links

In railway systems, the railway links are usually very long and easily exposed to flood events, especially in large flat geographical terrains. As railway stations in cities are usually well-built and can therefore typically resist flood events, this study focuses on the flood-induced vulnerability of railway links.

When a flood event occurs, the disruption of a railway link depends not only on the event characteristics, but also its own properties such as the geographical terrain, the year it is built, the maintenance conditions, etc. As such detailed physical property data of the railway links is unavailable for this study, it uses historical flood-induced railway link disruption event data and historical flood event data to compute the CRSN links’ vulnerabilities. The link vulnerability computation method is based on the perspective that if a railway link is located in a high frequency flood event area but has few instances of flood-induced disruption in the past, then its flood-induced vulnerability should be low. By contrast, if a link is located in a low frequency flood event area but has been disrupted by floods frequently in the past, then its vulnerability should be high.

Let $p_l$ ($0 \leq p_l < 1$) denote the flood-induced vulnerability of link $l$, $e_l$ the number of flood-induced railway link disruption occurrences on link $l$ in the past 30 years, and $f_l$ the number of flood events along link $l$ in the past 30 years. Then, the flood-induced vulnerability for link $l$ is:

$$p_l = \frac{e_l}{f_l}, l \in L.$$  

As only the data of flood disrupted railway links (but not the exact disrupted segment of a link) is available, two assumptions are made: (1) all segments along a railway link have the same failure probability; (2) different areas in a province have similar geographical and climatic conditions. Based on these two assumptions, the railway link failure probability under a flood is estimated based on the link length, the number of flood disruption events in the province, and the total length of railway links in the province. Similar to the component fragility under hurricanes and earthquakes, the component fragility under a flood is a function of the flood intensity. When detailed data is available such as flood depth, flood damage level, and component interruption frequency under a certain flood damage level, the component fragility can be modeled.
Then, $e_l$ can be computed as:

$$e_l = \frac{\text{(length of } l\text{)} \times \text{(number of disruption events in that province)}}{\text{total length of railway links in that province}}$$

As the spatial extent of each flood event to decide the set of affected railway links is not available in the study experiments, $f_l$ is set as the number of flood events in the province that includes link $l$. Note that when a railway link spans two or more provinces, $e_l$ and $f_l$ are computed for the associated segments in these provinces using GIS technology.

Based on the $p_l$ values, the 20 most vulnerable railway links in the CRSN are listed in Table 2 and illustrated in Fig. 4. Most of these links are located in three northeastern provinces that are located near the Songhua river basin where river flood events occur frequently. Another reason is that some of these railway links were originally built before 1949, with poorer quality and low design standards to resist flood events. The data in Table 2 is used in Section 3.2 for one of the vulnerability mitigation strategies (Strategy 2) compared in this study.

**Table 2** The 20 most vulnerable railway links in the CRSN.

<table>
<thead>
<tr>
<th>No.</th>
<th>Railway Link</th>
<th>$p_l$</th>
<th>No.</th>
<th>Railway Link</th>
<th>$p_l$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ShenYang-QinHuangDao</td>
<td>0.258</td>
<td>11</td>
<td>JiaoHe-AnTu</td>
<td>0.144</td>
</tr>
<tr>
<td>2</td>
<td>TuMen-MuDanJiang</td>
<td>0.196</td>
<td>12</td>
<td>QinHuangDao-Beijing</td>
<td>0.136</td>
</tr>
<tr>
<td>3</td>
<td>YanJi-JiaoHe</td>
<td>0.195</td>
<td>13</td>
<td>DuYun-HeChin</td>
<td>0.125</td>
</tr>
<tr>
<td>4</td>
<td>NanCha-SuiHua</td>
<td>0.183</td>
<td>14</td>
<td>NanCha-JiaMuSi</td>
<td>0.122</td>
</tr>
<tr>
<td>5</td>
<td>TongLiao-BaLinYou</td>
<td>0.163</td>
<td>15</td>
<td>NanNing-BaiCe</td>
<td>0.120</td>
</tr>
<tr>
<td>6</td>
<td>QingGang-HeiHe</td>
<td>0.162</td>
<td>16</td>
<td>LongHua-ChiFeng</td>
<td>0.118</td>
</tr>
<tr>
<td>7</td>
<td>MuDanJiang-JiXi</td>
<td>0.155</td>
<td>17</td>
<td>Changling-BaiCheng</td>
<td>0.115</td>
</tr>
<tr>
<td>8</td>
<td>SuiFenHe-MuDanJiang</td>
<td>0.154</td>
<td>18</td>
<td>NenJiang-LunChun</td>
<td>0.115</td>
</tr>
<tr>
<td>9</td>
<td>LongHua-ChengDe</td>
<td>0.149</td>
<td>19</td>
<td>ZhongWei-PingLiang</td>
<td>0.114</td>
</tr>
<tr>
<td>10</td>
<td>DanDong-Benxi</td>
<td>0.147</td>
<td>20</td>
<td>NanNing-ChongZuo</td>
<td>0.113</td>
</tr>
</tbody>
</table>
2.4 Computation of the CRSN vulnerability

Based on the flood event scenario generation method and the computed vulnerabilities of the railway links, this section introduces a model to compute system-level vulnerability of the CRSN. As introduced in section 1, this paper defines the vulnerability of a railway system as its performance drop under a given disruptive event. As trains have different timetables and the number of trains in the network is time-dependent, some performance metrics at hourly or smaller scale, such as the fraction of trains in operation at a specific time, are also time-dependent, and the corresponding vulnerability measures are time-dependent as well. Since the proposed flood-induced vulnerability assessment approach has a planning perspective, this paper uses a time-independent vulnerability metric which integrates the interruption durations. This vulnerability metric is the sum of the number of interrupted trains (which is defined as the drop of the number of trains which can run during a typical weekday) on each day of the whole interruption period. In addition, note that network congestion may occur after a disruptive event. Modeling network congestion requires more detailed information, such as the event occurrence time and location, each train’s location at the occurrence of the event, and so on. This paper simply assumes that all
interrupted trains can find some nearby stations or emergency railway segments to stop so that they would not cause congestion and affect other trains. Based on this assumption, the number of interrupted trains in one day of the interruption period can be computed using the railway link daily interruption status.

This study assesses the monthly CRSN vulnerability of June through September so as to mitigate its flood-induced vulnerability during flood season. The monthly flood-induced vulnerability of the CRSN is computed as the sum of daily vulnerability values in a month. The daily vulnerability value is the number of interrupted trains under a flood scenario, which depends on the status of each affected link under the flood scenario. Hence, one critical step is to determine the daily interruption status of each flood-affected railway link.

### 2.4.1 Determination of railway link daily interruption status

The following notation is used in the determination of the daily interruption status of each flood-affected railway link. Let \( X_l(d) = \{0, 1\} \) denote the interruption status of link \( l \) on the \( d^{th} \) day in the month of interest. If link \( l \) is interrupted on the \( d^{th} \) day, \( X_l(d) = 1 \); otherwise \( X_l(d) = 0 \). Let \( L^d \) denote the set of flood-affected links; they represent all links in the provinces subjected to flood events based on the scenario generated. Denote \( \tau_l \) as the interruption duration of link \( l \). Let \( L'(d) \) denote the set of interrupted railway links on the \( d^{th} \) day; that is, \( L'(d) = \{l \mid X_l(d) = 1, l \in L^d\} \). Let \( \delta_l(d) = \{0, 1\} \) denote the service status of the \( i^{th} \) train on the \( d^{th} \) day. If the \( i^{th} \) train is out of service on the \( d^{th} \) day due to a flood, \( \delta_l(d) = 1 \); otherwise \( \delta_l(d) = 0 \). Let \( L^p_l \) denote the set of railway links that the \( i^{th} \) train passes through; that is, \( L^p_l = \{l| r_{lt} = 1, r_{lt} \in R\} \) (matrix \( R \) is defined in Section 2.1).

A conditional daily updating process is proposed to determine the interruption status of each flood-affected railway link on each day. On the \( d^{th} \) day, if link \( l \) is interrupted under the flood scenario, then its interruption duration \( \tau_l \) is sampled from the probability distribution of the historical interruption duration time data. The values of \( X_l(d) \sim X_l(d + \tau_l - 1) \) will be set to be 1; that is, link \( l \) remains interrupted from the \( d^{th} \) day to \((d + \tau_l - 1)^{th}\) day; the status of link \( l \) on other days does not change. If link \( l \) is not interrupted under the flood scenario, the status of link \( l \) does not change for any day. The interruption status of link \( l \) from the \( d^{th} \) day to the last day of the month is determined through this process.

To illustrate the conditional daily updating process, an example is presented in Table 3. It shows the update of the interruption status of link \( l \) under flood scenarios over four days. Suppose link \( l \) is interrupted under the \( d^{th} \) day’s flood scenario, and the interruption duration is 4 days (\( \tau_l = 4 \)); that is, from the \( d^{th} \) day to the \((d+3)^{th}\) day. The values of \( X_l(d) \) through \( X_l(d+3) \) are set to 1, as
shown in the second row in Table 3. Suppose link \( l \) is interrupted under the \((d+1)\text{th}\) day’s flood scenario \((\tau_l = 5)\); the values of \(X_l(d+1)\) through \(X_l(d+5)\) are set to 1 (as highlighted by the bold italicized numbers in Table 3). Link \( l \) is not interrupted under the \((d+2)\text{th}\) day’s flood scenario \((\tau_l = 0)\). Hence, the interruption status of link \( l \) will not change; that is, the link interruption status for that day and the subsequent days is equal to the corresponding values on the previous day, as illustrated by the fourth row in Table 3. Similarly, if link \( l \) is interrupted under the \((d+3)\text{th}\) day’s flood scenario \((\tau_l = 2)\), the values of \(X_l(d+3)\) and \(X_l(d+4)\) are set to 1.

<table>
<thead>
<tr>
<th>Flood scenario</th>
<th>( \tau_l )</th>
<th>( X_l(d) )</th>
<th>( X_l(d + 1) )</th>
<th>( X_l(d + 2) )</th>
<th>( X_l(d + 3) )</th>
<th>( X_l(d + 4) )</th>
<th>( X_l(d + 5) )</th>
<th>( X_l(d + 6) )</th>
<th>( X_l(d + 7) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( d^{\text{th}} ) day</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>((d + 1)^{\text{th}} ) day</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>((d + 2)^{\text{th}} ) day</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>((d + 3)^{\text{th}} ) day</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### 2.4.2 Average CRSN vulnerability computation approach

The average monthly flood-induced vulnerability of the CRSN is computed using the Monte Carlo simulation procedure illustrated in Fig.5, summarized hereafter:

**Step 1:** Set the Monte Carlo iteration counter \( \alpha = 1 \).

**Step 2:** Initialize the interruption status profile of all links to 0 for the month of interest, that is, none of them are interrupted or \( X_l(d) = 0 \) for all \( l \) and \( d \). Set the start time \( d = 1 \), which is the first day of a month.

**Step 3:** Generate a daily flood scenario using the method described in Section 2.2 and determine the flood affected link set \( L^d(d) \) under this scenario.

**Step 4:** Determine the status of each flood affected railway link \( l \) from the \( d^{\text{th}} \) day of this month.

**Step 4.1:** For each railway link \( l \) in \( L^d(d) \), identify its status by comparing its flood-induced vulnerability \( p_l \) and a uniformly distributed random number \( b \) in \([0, 1]\). If \( p_l > b \), the railway link is assumed interrupted, that is, \( X_l(d) = 1 \), and go to Step 4.2 for generating the link’s interruption time; otherwise, go to Step 4.3.
Step 4.2: Generate the link’s interruption time $\tau_l$ using the probability distribution of the historical interruption duration time data, and set the railway link state $X_l$ from the $d^{th}$ day to $(d + \tau_l - 1)^{th}$ day as interrupted in this month, that is, $X_l(d) = 1, \ldots, X_l(d + \tau_l - 1) = 1$. This information represents the updated interruption status profile for link $l$ for the subsequent days of the month. Go to Step 4.3.

Step 4.3: Delete link $l$ from $L^A(d)$ and check whether $L^A(d)$ is empty. If $L^A(d)$ is empty, the status of all flood affected links is determined, go to Step 5; otherwise, repeat Step 4.
**Step 5:** Compute the CRSN vulnerability value on the $d^{th}$ day of this month.

Based on the definition of the system-level CRSN vulnerability above, the vulnerability of the CRSN on the $d^{th}$ day is equal to the number of interrupted trains on that day. The vulnerability of the CRSN on the $d^{th}$ day is computed as $\sum_{t=1}^{4196} \delta_t(d)$, where $\delta_t(d) = 1 - \prod_{i \in E_t} (1 - X_i(d))$.

**Step 6:** If the value of $d$ indicates that it is the last day of this month, compute the value of the CRSN vulnerability in this month as the summation of the corresponding daily values, and go to Step 7. Otherwise, update the day counter, $d = d + 1$, and repeat Steps 3-6.

**Step 7:** Update the Monte Carlo iteration counter. If $\alpha$ is less than 1,000,000, $\alpha = \alpha + 1$ and go to Step 2. If it is equal to 1,000,000, compute the average value of the CRSN vulnerability in this month over the 1,000,000 Monte Carlo simulations.

### 3. Numerical analysis and results

This section first applies the approach illustrated in Fig. 5 to determine the average flood-induced vulnerability of the CRSN from June through September. Second, to analyze whether the proposed vulnerability approach can better characterize the impact of flood events on the railway system, CRS flood-induced railway damage data for a few years and the currently-used metrics to indicate flood impact are compared with vulnerability values computed using the proposed approach in Fig. 5. Finally, as a simple application of the proposed vulnerability assessment approach, four vulnerability mitigation strategies are proposed and evaluated to compare their effectiveness.

#### 3.1 Monthly flood-induced CRSN vulnerability

The vulnerability of the CRSN for each of the four months, June through September, is computed using the methodology proposed in Section 2. The average monthly flood-induced vulnerabilities of the CRSN computed using the approach in Fig. 5 (using the 30-year historical data to generate the relevant probability distributions) are shown in Table 4. The results show that July and August have larger vulnerability values compared to other months; this is consistent with the historical flood event data in Fig. 2. The average vulnerability of the CRSN in the flood season (June-September) is 6775.1. It can be used to estimate the average impact of the flood season on the CRSN, and can be useful for planning and decision-making processes related to reducing the flood-induced railway damage and its impact.

<table>
<thead>
<tr>
<th>Month</th>
<th>June</th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>Flood Season</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average vulnerability value</td>
<td>1371.1</td>
<td>2456.8</td>
<td>2292.2</td>
<td>655.0</td>
<td>6775.1</td>
</tr>
</tbody>
</table>
Next, data for the flood season for a few years is used to analyze whether the Monte Carlo simulation based vulnerability assessment methodology illustrated through the vulnerability values in Table 3 is more representative of the impact/damage due to the flood events compared to currently-used indicators [1] in the CRS context. Data for the June-September months from four specific years (1991, 1995, 1996, and 1998) is used for comparison purposes, as only they have enough detailed disruption event data needed for the computation of the flood-induced vulnerability value, including the number of these flood events, and the specific interrupted railway links and their interruption duration under each event. Table 5 shows the flood-induced vulnerabilities computed for the June-September periods of these four years using the vulnerability computation approach discussed in Section 2.4, along with two currently-used metrics (flood events number and total link disruption duration) for the impact of floods on the CRS.

<table>
<thead>
<tr>
<th>Year</th>
<th>Flood events number</th>
<th>Total disruption duration</th>
<th>Flood-induced vulnerability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>9</td>
<td>91 link-days</td>
<td>3278</td>
</tr>
<tr>
<td>1995</td>
<td>5</td>
<td>110 link-days</td>
<td>9519</td>
</tr>
<tr>
<td>1996</td>
<td>21</td>
<td>278 link-days</td>
<td>8614</td>
</tr>
<tr>
<td>1998</td>
<td>12</td>
<td>268 link-days</td>
<td>9045</td>
</tr>
</tbody>
</table>

In many official documents related to the impact of hazards on the CRS, the number of hazard events and the total disruption duration in a year are usually used to describe the impact to the CRS under various natural or manmade hazards in that year [1]. A comparison of the three metrics (damage events number, total disruption duration and flood-induced vulnerability) in Table 4 illustrates a lack of consistency among them, leading to the question of which of them is most representative of the flood impact. The number of flood events and the total disruption duration in 1995 are much less than those in 1996 and 1998, but the flood-induced vulnerability in 1995 is higher than that of 1996 and 1998. The flood events number in 1995 is about half that of 1991, and the total disruption duration in 1995 exceeds that of 1991 by only about 20%, but the vulnerability value in 1995 is about 3 times that of 1991. History indicates that nationwide floods occurred in China in 1995 and 1998, and caused serious economic losses. Hence, the number of hazard events and the total disruption time may not be consistent indicators of the extent of flood impact on the CRS, especially in terms of the service disruption in the system. The proposed vulnerability metric, which includes the consequences of the flood damage to the infrastructure on the train service levels, can better quantify the CRS impact level under hazards.
3.2 CRS vulnerability mitigation strategies

Besides infrastructure vulnerability assessment, vulnerability mitigation strategies are also very important to infrastructure protection [61-63]. To mitigate the flood-induced CRS vulnerability, the Chinese government needs to spend a large amount of money to maintain the railway links before the flood season every year. The aim of the maintenance is to decrease the failure probability of the railway links under flood events. As a result of the monetary investment, the vulnerability of the railway link $p_l$ in Section 2.3 will be decreased after the maintenance. In this paper, we use the terms maintenance strategy and vulnerability mitigation strategy interchangeably.”

The maintenance strategy considers three factors: (1) which links should be prioritized for maintenance? (2) how many links should be simultaneously considered for maintenance? (3) to what extent should the maintenance mitigate the link vulnerability? Here, the maintenance is an abstract notion which may represent a certain maintenance action in the railway system, such as strengthening the soil foundation, replacing the sleeper and so on. The effect of maintenance is reflected by the reduced link vulnerability and the mitigation percentage in the link vulnerability is labeled “maintenance intensity” in this study. In this section, four maintenance strategies are introduced and their effects are compared in terms of one or more of the aforementioned factors.

Note that the maintenance strategy analysis here is mainly used to illustrate a simple application of the proposed vulnerability assessment approach. Due to the lack of maintenance cost data, this paper assumes that the maintenance cost of different railway links with various lengths is identical and does not consider the maintenance costs. When sufficient data is available, a future research direction is to integrate railway link length and maintenance cost into the maintenance strategies and then performs a comprehensive effectiveness analysis.

3.2.1 Flood-induced railway vulnerability mitigation strategies

Let $x$ denote the number of maintained links $x \in [0, 500]$, and $y$ denote the maintenance intensities which can range from 0 to 100%. The updated vulnerability of a railway link $l$ after maintenance is: $p_l^{new} = (1 - y) \times p_l^{old}$. For example, $y = 1$ means the corresponding railway link has an updated vulnerability zero, and will be not disrupted by flood events.

The first strategy RMS$(x,y)$ is random maintenance; in it, the railway links are chosen randomly for maintenance. $x$ railway links are chosen from the CRSN randomly, and their vulnerabilities are updated $p_l^{new} = (1 - y) \times p_l^{old}$. The flood-induced CRSN vulnerability is determined based on the updated link vulnerabilities $P = \{p_l^{new}\}$. 
The second strategy $FMS(x, y)$ chooses the railway links for maintenance based on their link vulnerabilities (see Table 2). The higher vulnerability links are maintained in priority by first sorting the railway links according to their vulnerability values and choosing the $x$ links with the highest vulnerabilities. The vulnerability values for these $x$ links are updated, $p_l^{\text{new}} = (1 - y) \times p_l^{\text{old}}$, and the flood-induced CRSN vulnerability is determined using the updated values.

The third strategy $BMS(x, y)$ chooses railway links for maintenance according to their importance for service, in terms of the number of trains traveling through a link (see Table 6). The railway links are sorted according to the number of trains traveling through them, and the top $x$ links are selected for maintenance. The vulnerability values for these $x$ links are updated and the flood-induced CRSN vulnerability is determined.

The number of trains passing through a railway link $l$ can be obtained using the matrix $R$ defined in Section 2.1, as $\sum_{t=1}^{196} r_{lt}$. The top 20 such links are listed in Table 6. They are the busiest links in the CRSN. Most of these links are located in the east and the south of China, such as Jiangsu and Guangdong provinces which are China’s richest provinces.

<table>
<thead>
<tr>
<th>No.</th>
<th>Railway Link</th>
<th># of Trains</th>
<th>No.</th>
<th>Railway Link</th>
<th># of Trains</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ShangHai-KunShan</td>
<td>427</td>
<td>11</td>
<td>ZhuZhou-ChangSha</td>
<td>327</td>
</tr>
<tr>
<td>2</td>
<td>Suzhou-KunShan</td>
<td>423</td>
<td>12</td>
<td>Shenzhen-DongGuan</td>
<td>312</td>
</tr>
<tr>
<td>3</td>
<td>WuXi-SuZhou</td>
<td>403</td>
<td>13</td>
<td>Guangzhou-DongGuan</td>
<td>302</td>
</tr>
<tr>
<td>4</td>
<td>WuXi-ChangZhou</td>
<td>395</td>
<td>14</td>
<td>TangShan-QinHuangDao</td>
<td>276</td>
</tr>
<tr>
<td>5</td>
<td>NanJing-ZhenJiang</td>
<td>395</td>
<td>15</td>
<td>Jinzhou-HuLuDao</td>
<td>268</td>
</tr>
<tr>
<td>6</td>
<td>ChangZhou-ZhenJiang</td>
<td>393</td>
<td>16</td>
<td>ShanHaiGuan-HuLuDao</td>
<td>262</td>
</tr>
<tr>
<td>7</td>
<td>HengYang-ZhuZhou</td>
<td>354</td>
<td>17</td>
<td>NanChang-JiuJiang</td>
<td>252</td>
</tr>
<tr>
<td>8</td>
<td>ShaoGuan-ChenZhou</td>
<td>349</td>
<td>18</td>
<td>NanChangXian-NanChang</td>
<td>251</td>
</tr>
<tr>
<td>9</td>
<td>ShaoGuan-GuangZhou</td>
<td>349</td>
<td>19</td>
<td>Tianjin-BeiJing</td>
<td>251</td>
</tr>
<tr>
<td>10</td>
<td>HengYang-ChenZhou</td>
<td>347</td>
<td>20</td>
<td>YingTan-NanChangXian</td>
<td>250</td>
</tr>
</tbody>
</table>

The distribution of railway links in terms of the number of trains traveling through various links is shown in Fig. 6. It illustrates the significant variability in the CRSN in this context; more than 270 railway links belong to $(0, 100]$ in terms of the number of trains traveling through them, but only a few links have a value exceeding 390.
The fourth strategy $FBMS(x, y)$ chooses links according to the product of their vulnerability and the number of trains that travel through them. This product can be viewed as a proxy for the risk associated with the link relative to impact under floods. Based on this strategy, the links are sorted according to this product, and the top $x$ links are chosen. The vulnerability values for these $x$ links are updated and the CRSN vulnerability is determined. The top 20 high risk railway links in the CRSN are listed in Table 7.

Table 7 The top 20 high risk railway links.

<table>
<thead>
<tr>
<th>No.</th>
<th>Railway Line</th>
<th>Value</th>
<th>No.</th>
<th>Railway Line</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TangShan-QinHuangDao</td>
<td>18.10</td>
<td>11</td>
<td>ZhangJiaKou-BeiJing</td>
<td>8.299</td>
</tr>
<tr>
<td>2</td>
<td>NanJing-BengBu</td>
<td>10.43</td>
<td>12</td>
<td>XinMin-JinZhou</td>
<td>7.820</td>
</tr>
<tr>
<td>3</td>
<td>JinZhou-HuLuDao</td>
<td>10.21</td>
<td>13</td>
<td>YingKou-WaFangDian</td>
<td>7.391</td>
</tr>
<tr>
<td>4</td>
<td>XuZhou-BengBu</td>
<td>10.03</td>
<td>14</td>
<td>TianJin-CangZhou</td>
<td>6.474</td>
</tr>
<tr>
<td>5</td>
<td>PuLanDian-DaLian</td>
<td>9.536</td>
<td>15</td>
<td>FuYang-HaoZhou</td>
<td>6.274</td>
</tr>
<tr>
<td>6</td>
<td>ShiJiaZhuang-BaoDing</td>
<td>9.252</td>
<td>16</td>
<td>ShenYang-JinZhou</td>
<td>6.212</td>
</tr>
<tr>
<td>7</td>
<td>XiAn-Chencang</td>
<td>9.223</td>
<td>17</td>
<td>BaZhou-HengShui</td>
<td>6.191</td>
</tr>
<tr>
<td>8</td>
<td>BeiJing-BaoDing</td>
<td>8.669</td>
<td>18</td>
<td>ShangHai-HangZhou</td>
<td>6.100</td>
</tr>
<tr>
<td>9</td>
<td>XingTai-ShiJiaZhuang</td>
<td>8.537</td>
<td>19</td>
<td>DanDong-BenXi</td>
<td>5.881</td>
</tr>
<tr>
<td>10</td>
<td>NanCha-SuiHua</td>
<td>8.420</td>
<td>20</td>
<td>LouDi-HuaiHua</td>
<td>5.868</td>
</tr>
</tbody>
</table>

### 3.2.2 Analyzing the effectiveness of the vulnerability mitigation strategies

In this section, the CRSN vulnerability is computed using the four maintenance strategies and used to compare the effectiveness of these strategies for different number of links to be maintained, $x$. Then, the effect of different maintenance intensities on the FBMS strategy is analyzed.
The effectiveness of the four strategies under different number of maintenance links is analyzed by varying the value of \( x \) from 1 to 500, while setting \( y \) to 1. Fig. 7 illustrates the performance of the four strategies for the month of July (similar trends are observed for the other three months). Each point in Fig. 7 represents the average CRSN vulnerability value over 1,000,000 Monte Carlo simulations. With the increase in the number of maintained links, the CRSN vulnerability value decreases for all strategies. Also, as illustrated by the figure, the most effective maintenance strategy is FBMS. This is because the FBMS not only considers the flood impact on the railway link itself (link vulnerability), but also factors the link’s influence on the whole railway system (number of trains that use the link). The least effective strategy is to select railway links randomly for maintenance as this strategy is not related to system/flood characteristics. The effectiveness of the FMS and BMS strategies is close to each other when the number of maintained links is below 200. When the number of maintained links is above 200, FMS does better than BMS. When the number of maintained railway links is above 350, the vulnerability curves of FMS and FBMS are close to zero. This is because the number of links whose vulnerability is above 0 is 375. After these 375 links are maintained, no railway link disruption event will occur under flood events as all links have vulnerability zero, and therefore the CRS vulnerability value also equals zero.
The effect of maintenance intensity is analyzed using the FBMS strategy. High maintenance intensity usually requires more resources. When resources are limited, determining the maintenance intensity and the number of railway links to be maintained is complex due to the large number of potential combinations to consider. There are two choices, a higher maintenance intensity and fewer maintained links, or a lower maintenance intensity and several maintained links. Here, the effect of seven levels of maintenance intensities is compared for the month of July, as shown in Fig. 8. The maintenance intensity $y$ is set to 100%, 80%, 60%, 50%, 30%, and 10%, respectively, and the number of maintained links $x$ is varied from 1 to 400. As the maintenance intensity increases, the average CRSN vulnerability curve lowers sharply. When the maintenance intensity is low, such as 10% and 30%, the effect of the maintenance is not significant. This suggests that if the maintenance intensity is low, increasing the number of maintenance links has little effect on the vulnerability mitigation. Hence, decision-makers should consider increasing the maintenance intensity rather than focusing on just the number of maintained links in such scenarios. Based on the insights from Figs. 7 and 8, an effective maintenance plan is to distribute the limited resources to maintain the high risk railway links in the FBMS strategy and ensure the non-interruption of these links under the flood events.

4 Conclusions

China is a country prone to floods. Its railway system plays a critical role in its economy. Analyzing how to assess the influence of flood events on the railway system and identifying an effective maintenance strategy are important to the CRS protection and its effective functioning. This paper proposes a methodology to quantitatively assess the vulnerability of the CRS under flood events using available historical data and GIS technology. In it, flood event scenarios are
first generated based on flood event data for the past 30 years. This data, along with the flood-induced railway disruption event data, is used to compute the railway link vulnerabilities. Based on these, a system-level average vulnerability is determined for the proposed network representation of the CRS. Numerical analyses illustrate that the proposed system-level vulnerability metric is more representative of the flood impact on the railway system compared to other indicators currently used (such as number of flood events, and the disruption duration). Four vulnerability mitigation maintenance strategies are analyzed, and provide insights on the factors that should be considered for the effective maintenance of the CRS so as to minimize the impact of flood events. The proposed methodology can be used not only for flood events, but also for other types of hazards such as earthquakes and hurricanes.

In this study, based on the available data, the system-level vulnerability is determined using the number of interrupted trains and their interruption duration. If more specific data related to the number of affected passengers and their waiting times were available, the vulnerability metric could be directly linked to passenger service levels. Further, the proposed methodology could be improved by: (1) developing a more comprehensive flood scenario model and component fragility model using more detailed data collected from the field; (2) improving maintenance strategies by simultaneously considering railway link length and maintenance cost; and (3) introducing a time-dependent vulnerability assessment model for operational purposes. The consideration of maintenance resource limitations is another future research direction; however, this entails the availability of the total maintenance resource amount and per length railway link maintenance cost. The proposed methodology can also be extended to assess vulnerability of other transportation systems under threats from natural disasters.

**Acknowledgements**

This work is jointly supported by National Natural Science Foundations of China (60903174, 51208223, and 61433006), the Natural Science Foundations of Hubei province (2012FFB02203), the Fundamental Research Funds for the Central Universities, HUST: 2012QN087, 2012QN088, and the U.S. Department of Transportation through the NEXTRANS Center, the USDOT Region 5 University Transportation Center.

**Appendix A. List of variable**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>G = (S, L)</td>
<td>The network representation of Chinese railway system with node set S and arc set L, where S denotes the set of railway stations, and L denotes the set of railway links connecting these stations.</td>
</tr>
<tr>
<td>T = {t}</td>
<td>The set of trains</td>
</tr>
<tr>
<td>R = {r_{tl}}</td>
<td>If train t passes through the railway link l, r_{tl}=1, otherwise, r_{tl}=0, where t ∈ T, l ∈ L.</td>
</tr>
<tr>
<td>p_l</td>
<td>The flood-induced vulnerability of link l</td>
</tr>
<tr>
<td>p_{l}^{old}</td>
<td>The flood-induced vulnerability of a railway link l before maintenance</td>
</tr>
</tbody>
</table>
References


