How to use empirical data to improve transportation infrastructure risk assessment

Weihua Zhu1,2, Kai Liu1,2∗, Ming Wang1,2, Sadhana Nirandjan3, Elco E. Koks3

1 School of national security and emergency management, Beijing Normal University, Beijing 100875, China
2 Academy of Disaster Reduction and Emergency Management, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
3 Institute for Environmental Studies (IVM), Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands

Correspondence to Kai Liu (liukai@bnu.edu.cn).

Abstract:

Rainfall-induced hazards, such as landslides, debris flows, and floods cause significant damage to transportation infrastructure. However, an accurate assessment of rainfall-induced hazard risk to transportation infrastructure is limited by the lack of regional and asset-tailored vulnerability curves. This study aims to use multi-source empirical damage data to generate vulnerability curves and assess the risk of transportation infrastructure to rainfall-induced hazards. The methodology is exemplified through a case study for the Chinese national railway infrastructure. In doing so, regional and national-level vulnerability curves are derived based on historical railway damage records. This is combined with daily precipitation data and the railway infrastructure market value to estimate regional- and national-level risk. The results show large variations in the shape of the vulnerability curves across the different regions. The railway infrastructure in Northeast and Northwest China is more vulnerable to rainfall-induced hazards due to low protection standards. The expected annual damage (EAD) ranges from 1.88...
to 5.98 billion RMB for the Chinese railway infrastructure, with a mean value of 3.91 billion RMB. However, the risk of railway infrastructure in China shows high spatial differences due to the spatially uneven precipitation characteristics, exposure distribution, and vulnerability curves. The South, East and Central provinces have a high risk to rainfall-induced hazards, resulting in an average EAD of 184 million RMB, 176 and 156 million RMB, respectively, whereas the risk in the Northeast and Northwest provinces are estimated to be relatively lower. The usage of multi-source empirical data offer opportunities to perform risk assessments that include spatial detail among regions. These risk assessments are highly needed in order to make effective decisions to make our infrastructure resilient.

Keywords: multi-source empirical data, vulnerability curve, risk estimate, damage length factor

1. Introduction

In recent years, extreme precipitation events have increased in both frequency and intensity in the context of global warming (Shi et al., 2018; Cardoso Pereira et al., 2020; Li et al., 2020). Extreme precipitation may generate landslides, debris flows, and floods, which have the potential to damage transportation infrastructure and disrupt transportation functions, thereby posing a severe threat to the economy and society (Pregnolato et al., 2017; Diakakis et al., 2020; Petrova, 2020). In July 2021, Zhengzhou was hit by a heavy downpour, that reached a cumulative precipitation of 617.1mm in three days. The associated flash floods resulted in the destruction of the Zhengzhou metro system; suspension of more than 80 bus lines; damage to 67 urban bridges, culverts and tunnels; cancellation and delay of more than 200 flights from
Zhengzhou airport; and flooded lines, collapsed roadbeds, and waterlogging of equipment forcing railway operators to shut down for several days (Fig. 1a). On May 23, 2010, a landslide occurred in the Yujiang-Dongxiang section of the Shanghai-Kunming Railway in Jiangxi Province, causing the derailment of passenger train K859 (Fig. 1b). The sliding body was 60 metres long, 30 metres wide, and 3-8 metres thick, resulting in a volume of approximately 9,000 cubic metres. The cumulative precipitation in the 11 days before the incident was 251.5 mm in Xiaogang town, Dongxiang County. In China, the average annual direct damage of railway infrastructure caused by rainfall-induced hazards was approximately 3.29 billion RMB from 2000 to 2017 and has increased in recent years (Editorial Board of China Railway Yearbook, 2001-2017).

Fig. 1 (a) Transportation infrastructure damaged by floods triggered by extreme precipitation at Zhengzhou, Henan province (2021); (b) Railroad damage by a debris flow triggered by extreme precipitation at Xiaogang, Jiangxi Province (2010).

Accurate assessment of transportation infrastructure damage and risk due to hazards triggered by rainfall is an essential component in transportation infrastructure risk management (Liu et al., 2018a, 2021). In general, transportation infrastructure impacts due to natural hazards include two aspects: (1) direct damage to the structure (Koseki et al., 2012; Kellermann et al., 2015; Koks et al., 2019); and (2) indirect impact to the transportation service and associated
macroeconomic impact (Lamb et al., 2019). Determining direct damage is commonly done using vulnerability curves (Englhardt et al., 2019; Koks et al., 2019), which typically present the damage degree of infrastructure assets that would occur at specific hazard intensities (Jongman et al., 2012; Ward et al., 2013). As the critical link of hazard characteristics and damage loss, few studies (e.g. Sande and C.J, 2001; Kok et al., 2004; Huizinga et al., 2017) tried to work on vulnerability curves for transportation infrastructure assets in different regions. In these studies, empirical and synthetic approaches are usually adopted to develop curves based on damage data (Merz et al., 2010) and expert judgement (Gerl et al., 2016). Unfortunately, due to the lack of detailed damage data, such damage curves are unavailable for most regions. Habermann and Hedel (2018) conducted a literature review on the damage functions for transportation infrastructure due to wildfires and floods. They found that damage functions for the transportation sector are scarce in the literature, and damage curves for the transportation sector in different publications vary in shapes and values.

This article aims to use multi-source empirical damage data to assess the vulnerability and risk of transportation infrastructure associated with rainfall-induced hazards (i.e. landslides, debris flows, and floods). We develop a first set of regional and national vulnerability curves for Chinese railway infrastructure that relates the damage degree of railway assets to precipitation intensities. Based on these vulnerability curves, the risk of the railway infrastructure associated with rainfall-induced hazards is estimated.

The remainder of the article is organized as follows. Section 2 describes this work’s datasets, including data on precipitation, historical railway damage, yearly railway damage and railway
market value. Section 3 describes the methodological framework, thereby elaborating on the method for: (1) vulnerability assessment, (2) and risk estimation. Section 4 presents the main results. Sections 5 and 6 discuss the results and conclude the article.

2. Data collection

2.1 Precipitation data

The CN05.1 dataset provides information on the observed daily precipitation from 1961 to 2018 at a 0.25° spatial resolution (Wu and Gao, 2013; Zhang et al., 2019). The dataset is derived from more than 2,400 in situ gauging stations by the Chinese Meteorological Administration (CMA). The CN05.1 product has been recognized to be more reliable than its previous versions because of the inclusion of more ground stations (Yatagai et al., 2009; Zhang et al., 2019). The resolution of the CN05.1, however, is too coarse to accurately capture local extreme precipitation events. As a complementary precipitation dataset, we therefore extract local precipitation information from multiple news sources for 37% of the damage records (see section 2.2). These news sources contain precipitation data obtained from rain gauges installed by the railway department, thus measuring local extreme precipitation.
2.2 Historical railway damage by rainfall-induced hazards

Fig. 2 (a) Spatial distribution of national railway damage records. We divide the mainland of China into seven geographical divisions: Central China (I), East China (II), North China (III), Northeast China (IV), Northwest China (V), South China (VI), and Southwest China (VII).

(b) Temporal distribution of historical damage to the national railway infrastructure by rainfall-induced hazards from 2000 to 2016. Railway geometries © OpenStreetMap contributors 2019. Distributed under the Open Data Commons Open Database License (ODbL) v1.0.

Zhao et al. (2020) catalogued 464 railway disasters caused by rainfall-induced hazards in the Chinese railway system between 2000 and 2016. After removing service disruption disasters (i.e. trains that slow down or stop for safety reasons) that are irrelevant for this study, we found a total of 236 railway damage records that represent structural damage to railway assets or debris covering the rail. The spatial distribution of the filtered set of national railway damage records is presented in Fig. 2a. For all these records, we collect information about the occurrence date of the damage, the damage location, and the descriptive damage state by using online publicly available news sources. In this study, if damages induced by a precipitation event occurred in the segment between two adjacent stations, one damage record is counted.
Table 1 gives some typical railway damage records over 1981–2016. The information of the all damage records used in this study can be found in supplement material.

| Damage date | Url | Railway name | Damag segment | Damage state |
|-------------|-----|--------------|---------------|--------------|
| 2005/6/21   | http://news.sina.com.cn/c/2005-06-25/10586266797.shtml | Yingxia railway | Panfang-Yangkou | Geological, severe |
| 2005/6/21   | http://news.sina.com.cn/c/2005-06-25/10586266797.shtml | Yingxia railway | Xiawangtang-Shaikou | Geological, severe |
| 2013/7/13   | http://www.eeb.cn/tabid/372/infoid/1521/frtid/89/default.aspx | Baoxi railway | Yanan-Yananbei | Embankment, Moderate |
| 2016/7/17   | https://baike.so.com/doc/24425372-25257771.html | Jiaoliu railway | Wanyan-Longbizui | Track, severe |

For the available 236 railway damage records, 84% occurred in the summer (June, July, and August). Most of the disasters occurred in July, accounting for 40% of the 236 railway damage records; 30% and 14% occurred in June and August, respectively. These numbers correspond to most parts of China's rainy seasons in which precipitation is a crucial trigger of rainfall-induced hazards. Fig. 2a shows the spatial distribution of railway damage for the years 2000-2016. The results show that the national railway lines suffered widespread rainfall-induced damage, especially in South China. Detailed spatial distributions of damages and associated reasons were explored in previous research of Liu et al. (2018) and Zhao et al. (2020). To explore the spatial distribution of railway vulnerability in different regions, China was divided into seven sub-regions based on seven geographical divisions (Liu et al., 2020), as shown in Fig. 2a:
2.3 Railway damage yearly data

Two datasets are used to obtain railway damage yearly data: the national railway yearbooks (Editorial Board of China Railway Yearbook, 2001-2017) and the Zhengzhou regional administrator’s yearbooks (Editorial Board of Zhengzhou Administrator’s Railway Yearbook, 2001-2017). The national railway yearbooks cover data on the direct damage, total damage length and the number of total damage events (one damage event is defined as a main railway line is damaged by a precipitation event) per year for the national railway system. The Zhengzhou regional administrator’s yearbooks provide information on the number of total damage events and the total number of damaged places (i.e., a continuous section of damage) per year for the Zhengzhou administrator railway system (ZHR), from 2000 to 2017 by rainfall-induced hazards.

An overview of the yearly railway damage obtained from the two sources is shown in Table 2. Fig. 2b shows the direct damage per year from 2000 to 2017 (missing data in 2003 and 2004); the economic damage significantly increased from 2000 to 2017, which is due to the increased railway exposure and extreme precipitation events (Zhao et al., 2020). The average annual economic damage is estimated to be 3.29 billion RMB. The ZHR damage data shows that each damage event causes multiple damage places on railway infrastructure, with an average of nine damage places per event. Assuming that the number is the same for the national railway system, we calculate that the average damage length is 753 m per damage place for an event using the total number of damage events and total damage length at a national level.
Table 2 Railway damage for the period 2000-2017

| Year | National | Zhengzhou Administrator |
|------|----------|-------------------------|
|      | Damage event times | Damage length (km) | Direct damage (billion) | Damage event times | Damage places |
| 2000 | 183      | 478.6                   | 1.179                   |                   |               |
| 2001 | 98       | 358.8                   | 1.266                   | 42                | 469          |
| 2002 | 106      | 441.1                   | 1.156                   | 21                | 174          |
| 2003 | 142      |                        |                         | 81                | 125          |
| 2004 | 122      |                        |                         | 114               | 224          |
| 2005 | 203      | 2.105                   |                         | 36                | 169          |
| 2006 | 128      | 922.8                   | 2.073                   | 34                | 170          |
| 2007 | 121      | 832.6                   | 2.081                   | 40                | 247          |
| 2008 | 75       | 802.1                   | 1.911                   | 7                 | 272          |
| 2009 | 86       | 511.1                   | 1.741                   | 17                | 226          |
| 2010 | 177      | 1066.6                  | 6.473                   | 20                | 354          |
| 2011 | 109      | 1107.0                  | 2.767                   | 8                 | 144          |
| 2012 | 99       | 1606.0                  | 4.833                   | 41                | 160          |
| 2013 | 113      | 709.0                   | 6.280                   | 65                | 144          |
| 2014 | 82       | 654.0                   | 4.774                   | 52                | 206          |
| 2015 | 91       | 265.0                   | 3.576                   | 37                | 90           |
| 2016 | 211      | 388.0                   | 5.923                   | 53                | 246          |
| 2017 | 165      | 488.0                   | 4.531                   | 37                | 205          |

2.4 Railway market value

The railway market value is from the World Bank Office, China (Gerald Ollivier, 2014). They provide the range of average unit costs for the 200 km/h double-track railway (AUC-200D) shown in Table 3. The AUC-200D divides the cost of the railway into five elements, (1)
land acquisition and resettlement, and four first-level structures: (2) civil works (embankment, bridge or trunk), (3) track, (4) signalling, and (5) communications and electrifications. We use the mean value to present the unit cost of the element (e.g. the mean value is 6 million/per km of track element). The average railway market value used in this work is 56 million RMB, which does not consider land acquisition and resettlement costs since those parts were paid before construction.

### Table 3 Range of average unit costs (RMB million/per km of double track)

| Element                        | RMB million/per km of double track | Average unit costs (RMB million/per km of double track) |
|--------------------------------|------------------------------------|--------------------------------------------------------|
| Land acquisition and resettlement | 5-8                               | 6.5                                                    |
| Civil Works                    |                                    |                                                        |
| Embankment                     | 23-28                              |                                                        |
| Bridges/viaducts               | 59-62                              | 42.5                                                   |
| Tunnels                        | 51-68                              |                                                        |
| Track (ballasted)              | 5-7                                | 6                                                      |
| Signalling and communications  | 3-4                                | 3.5                                                    |
| Electrification                | 4                                  | 4                                                      |

### 3. Methods

Figure 3 presents an overview of the methodological framework used in this study. The methods in this study are divided into two parts: (1) vulnerability assessment and (2) risk estimation. In the first part, national and regional vulnerability curves that characterize the railway susceptibility by relating the damage degree to precipitation intensity are generated. In
the second part of the research, we estimate the risk to the Chinese railway system. The railway market value is combined with the vulnerability curve developed in the first part of the research and spatial data on the precipitation intensity to calculate the risk represented by expected annual damage (EAD).

Fig. 3 Methodology of using the multiple sources of data to estimate vulnerability and risk. Railway geometries © OpenStreetMap contributors 2019. Distributed under the Open Data Commons Open Database License (ODbL) v1.0.

3.1 Vulnerability curve estimation

3.1.1 Precipitation intensity estimated for damage records

The 88 damage records that are provided with additional local precipitation information from the news are shown in Fig. 4a with red lines. For each remaining damage record, we use the maximum 1-day precipitation amount along the damaged segment in the five consecutive days (M1-5d) before the damage occurred to present the precipitation intensity, shown in Fig. 4a.
with black lines. To keep the consistency of the precipitation, we use the extracted precipitation information from the news to correct the M1-5d. The relationship between precipitation from news and M1-5d is given in Eq. (1) and derived using a least-squares fitting method, as presented in Fig. 4b, with R square 0.63. The constructed curve allows us to transform the precipitation in CN05.1(pre(CN05.1)) to the local precipitation as far as possible.

$$pre(\text{news}) = 1.87 \times pre(\text{CN05.1}) + 27.35$$  \hspace{1cm} (1)

Fig. 4 (a) Spatial distribution of precipitation extracted from news and CN05.1; (b) The relationship between precipitation extracted from news and CN05.1.

3.1.2 Calculation of the damage ratio

The damage ratio is the ratio of the cost of repairing to the cost of rebuilding (Mazzorana et al., 2009), which is estimated by the news information and AUC-200D. First, we generate a custom damage ratio table based on the AUC-200D and the descriptive damage state is given in Sections 2.2 and 2.4. Second, we transform the descriptive damage state into a numerical damage ratio using the damage ratio table. There are three steps that we use to build the custom damage ratio table:
1. Determine the cost ratio of the railway value for first-level structures. Based on AUC-200D in Table 3, we calculate the cost ratio among four first-level structures. Taking the civil works: embankment as an example, the ‘total cost’ ratio is 0.65, 0.25, 0.09, 0.1 for embankment track, signalling, communications and electrifications, respectively. For bridges and tunnels, the total cost ratios are shown in Appendix Table A2.

2. Classify the damage state for each first-level structure and give the numerical damage ratio range for each classification. The final damage states and the associated numerical ratio are divided into four classifications, namely, total damage (1), severe damage (0.66-0.99), moderate damage (0.33-0.66), and light damage (0.01-0.33), as shown in Table 4.

3. Calculate the numerical damage ratio range for a combination of a railway structure and a damaged state, and determine associated damage descriptive information based on news sources for each combination. The final damage ratio table is presented in Table 4, which is multiplied by the cost ratio of a first-level structure and the range ratio of the damage state classification. We then classify the damage state in sec 2.2 into each category.

Based on the damage ratio table and the historical news, we obtain the numerical damage ratio for each record. For each event, three damage ratios, namely, minimum ratio, average ratio, and maximum ratio, are obtained based on the damage ratio range. For example, the minimum ratio for the embankment's severe damage state is 0.644, and the average and maximum ratios
are 0.536 and 0.429, respectively.

Table 4 Damage ratio table

| Element                  | Unit cost ratio | Damage state | Damage ratio | Description                                                                 |
|--------------------------|-----------------|--------------|--------------|-----------------------------------------------------------------------------|
| Embankment               | 0.6500          | Total        | 0.6500       | Total damage                                                                |
|                          |                 | Severe       | 0.4290-0.6435| Suspended sleepers; Hanging rails                                           |
|                          |                 | Moderate     | 0.2145-0.4290| Subgrade shoulder, drainage ditch, side drain, revetment slope protection, protecting wall: moderate damage, collapse |
|                          |                 | Slight       | 0.0065-0.2145| Subgrade shoulder, drainage ditch, side drain, revetment slope protection, protecting wall: mild damage, cracks, blockage, loose, wash out |
| Track                    | 0.1500          | Total        | 0.1500       | Total damage                                                                |
|                          |                 | Severe       | 0.0990-0.1485| Near-failure of components: sleepers, rail, track bed                        |
|                          |                 | Moderate     | 0.0495-0.0990| Two-component failure: sleepers, rail, track bed                             |
|                          |                 | Slight       | 0.0015-0.4950| Single-component failure: sleepers, rail, track bed                          |
| Signalling and communications | 0.0900        | Total        | 0.0900       | Total damage                                                                |
|                          |                 | Severe       | 0.0594-0.0891| Near-destruction of components: digital tuning and TDCS equipment           |
|                          |                 | Moderate     | 0.0297-0.0594| One-component destruction: digital tuning and TDCS equipment                 |
### 3.1.3 Fitting the vulnerability curves

We choose the log-normal distribution to fit the vulnerability curve. The cumulative distribution function of log-normal distribution is shown in Eq. 2,

\[ P(x) = \Phi\left[ \frac{\ln(x/\phi)}{\xi} \right] \]  

(2)

which has two parameters, the location parameter \( \phi \) and the scale parameter \( \xi \), namely, the median and standard values, respectively (Porter et al., 2007). We use the precipitation intensity as the \( x \) value and the damage ratio as the \( P(x) \) value. A log-normal vulnerability function is chosen because it is a parsimonious two-parameter distribution with positive support (ensuring that unrealistic negative loads cannot occur) and has many precedents for its use in fragility analysis (Porter et al., 2007).

In this study, we generate a total of seven vulnerability curves for the railway system: one for each of the six sub-regions (we combine North China into Central China since the damage records are less in North China), and one at the national level. To eliminate the noise and significant changes in the damage ratio, a moving average method is used to smooth the damage ratio in each precipitation intensity range. We use the criteria for classifying the precipitation intensity issued by the China Meteorological Administrator (2008), which is presented in Table...
Table 5 Classification of the precipitation intensity

| Precipitation intensity       | Total precipitation, in 24 h/mm |
|------------------------------|---------------------------------|
| Light rain                   | 0.1-9.9                         |
| Light rain-Moderate rain     | 5.0-16.9                        |
| Moderate rain                | 10.0-24.9                       |
| Moderate rain-Heavy rain     | 17.0-37.9                       |
| Heavy rain                   | 25.0-49.9                       |
| Heavy rain-Torrential rain   | 33.0-74.9                       |
| Torrential rain              | 50.0-99.9                       |
| Torrential rain-Downpour     | 75.0-174.9                      |
| Downpour                     | 100.0-249.9                     |
| Downpour-Heavy downpour      | 175.0-299.9                     |
| Heavy downpour               | ≥ 250.0                         |

3.2 Risk assessment

To calculate the direct risk to the Chinese railway infrastructure, we develop precipitation maps for different return periods based on the Gumbel distribution. From the daily precipitation time series in the CN05.1 product (1961-2018), we extract an annual time series of maximum precipitation volumes for 1961-2018. For each cell, we then fit a Gumbel distribution (Nadarajah, 2010) through this time series based on non-zero data. These Gumbel parameters are used to calculate precipitation volumes per grid cell for selected return periods (2, 5, 10, 25, 50, 100, 200, 250, 500, and 1000 years). Precipitation volumes are calculated conditionally on the exceedance probability of zero precipitation volume. For those cells where less than five non-zero data points are available, the precipitation volume is assumed to be zero (Ward et al.,
Risk is generally calculated by combining the hazard intensity, vulnerability, and exposure (Merz et al., 2009; Lamb et al., 2019). In this study, we present risk as expected annual damage (EAD) (Merz et al., 2009). The EAD is defined as the average expected yearly market loss and is estimated based on selected discrete hazard events with different return periods. The EAD is calculated using the trapezoidal rule (Espinet et al., 2018). The EAD of the Chinese railway system is expressed in Eq. 5 as follows:

$$\text{EAD} = \frac{1}{2} \sum_{r=1}^{n} \left( \frac{1}{T_r} - \frac{1}{T_{r+1}} \right) (D_1 + D_{r+1})$$

(3)

where $T_r$ is the $r^{th}$ return period, $D_i$ is associated with damage to the railway infrastructure, which is defined in Eqs. (4) and (5):

$$D_i = \Sigma_i^N H_i^r \ast V \ast E_i \ast C_{DL}$$

(4)

$$C_{DL} = \frac{D_L}{L}$$

(5)

where $H_i^r$ is the precipitation intensity amount of raster cell $i$ with a return period of T-year, $V$ is the vulnerability curve, $E_i$ is the railway market value of raster cell $i$, $N$ is the number of raster cells that intersect the railway line, and $C_{DL}$ is a damage length factor for calibration. In Eq. 5, $D_L$ is the average damage length (753 m) per damage place in an event, and $L$ is the average railway length for all raster cells that intersect with railway lines. This study and previous studies assume that assets exposed in one raster cell are exposed to the same damage degree for a certain hazard intensity. Based on the yearly railway damage data (sec 2.3), the average damage length in one damage place per event is 753 m. This is much shorter compared to the precipitation resolution (ca. 28 km) used in this work and is also shorter than the average railway length in each cell (ca. 14.6 km for double-track lines). We, therefore,
introduce a damage length factor ($C_{DL}$) to calibrate the estimated damage, assuming that not
the entire railway section in a specific cell suffers damage from an event.

4. Results

4.1 Vulnerability curves

The national- and regional-level vulnerability curves are presented in Fig. 5. The upper
boundary is the maximum vulnerability curve, the lower boundary is the minimum vulnerability
curve, and the middle black line is the average vulnerability curve, fitted by maximum,
minimum, and average ratios, respectively. Vulnerability curves have noticeable regional
differences across the country. When considering relatively low precipitation intensities,
railway lines in Northwest China are vulnerable to rainfall-induced hazards. Damage ratios in
Northwest China are higher than other regional- and national-level damage ratios with the same
precipitation intensity. For example, when the precipitation is 100 mm (torrential rain), the
national railway damage ratio is 0.124, whereas the railway damage ratio in Northwest China
is about 0.148. Railway lines in Northwest and Northeast China are particularly vulnerable to
rainfall events with high precipitation intensities. In case of extensive precipitation of more than
200 mm (downpour), the national railway damage ratio is approximately 0.175, the railway
damage ratio in Northeast China is about 0.180, and the railway damage ratio in Northwest
China can reach 0.212. In Northwest China, the precipitation amount over 100 years is less than
100 mm. Considering the low frequency of extreme precipitation and the expensive cost of high
protection standards, the Northwest China railway infrastructures are not robust relative to other
areas when looking at the same precipitation. In Northeast China, the oldest railway lines, which
have not been updated, have relatively low design standards and inadequate drainage facilities
to defend against extreme precipitation, resulting in higher vulnerabilities compared to other
regions.

Fig. 5 National and regional vulnerability curves between precipitation (mm) and damage
ratio. The maximum $R^2$ is the $R$ square for the maximum vulnerability curve, the average
$R^2$ is the $R$ square for the average vulnerability curve, the minimum $R^2$ is the $R$ square for
the minimum vulnerability curve.

4.2 Risk analysis

To incorporate the regional characteristics of the vulnerability for the Chinese railway system,
we use the regional vulnerability curves to assess the risk of the Chinese railway system. We
calculated the annual direct damage to railway infrastructure from 2000 to 2017, of which the
results are presented in Fig. 6. The grey area is the range of annual direct damage, with the
upper boundary calculated based on the maximum vulnerability curve, the lower boundary
calculated based on the minimum vulnerability curve, and the middle darker grey line
calculated based on the average vulnerability curve, thereby using the regional vulnerability
curves. The darker yellow dots are the annual statistical damage in the yearbook with a 10%
error scale. Compared to the statistical damage, we find that estimated damage is underestimated for high annual damage and overestimated for small annual damage, which is a consequence of using mean damage ratios for each precipitation range in the vulnerability curves. Damage in 2000, 2001 and 2002 is overestimated, and the estimated damage is calculated with the minimum vulnerability with 38%, 15% and 22% deviation from the statistical damage. For the left, 81.25% of the statistically damaged points are located in the estimated damage range. These results illustrate that the fitted vulnerability curves can be used to calculate the damage.

Fig. 6 Annual direct economic loss range
Damage in yearbook ± 10%

The regional and national EAD to railway infrastructure due to rainfall-induced hazards are presented in Fig. 7 using regional vulnerability curves. The national railway EAD is
approximately 3.91 billion RMB when calculated with average vulnerability curves. When calculated with minimum and maximum vulnerability curves, the national EAD is 1.88 billion RMB and 5.98 billion RMB, respectively. Regionally, damage among areas differs substantially. Using the results calculated with average vulnerability as an example, East China has the highest risk with approximately 1.0 billion RMB, which exceeds the national EAD with 25.5%. North, South and Southwest China face a similar risk, with approximately 15%, 14%, and 13% of total national damage, respectively. High-density railway infrastructure exposure combined with a high frequency of extreme precipitation in these regions results in railway infrastructure with the highest risk.

Fig. 7 Rainfall-induced hazard risk per region using different vulnerability curves. (a) The minimum vulnerability curve; (b) The average vulnerability curve; (c) The maximum vulnerability curve. The numbers on top of each stacked column chart are the national EAD values using the different vulnerability curves.

The national EAD per kilometre ranges from 32 to 86 thousand RMB, with an average of 65.38 thousand RMB using the average vulnerability curve. The EAD per kilometre is the highest in South China using the average vulnerability curves, which is 116.11 thousand RMB, followed by East China and Southwest China, for which the numbers are 96.30 and 87.37 thousand RMB, respectively. The railways in South, East and Southwest China require much
attention and must improve their robustness.

The risk per province calculated using the regional average vulnerability curves of the railway infrastructure to the rainfall-induced hazards are presented in Fig. 8. The risk differs considerably between regions when expressed in total EAD and EAD per kilometre. An examination of the total EAD shows that the provinces in North China, such as Hebei, Shanxi, Shandong, Henan, Southwest Sichuan and South Guangdong, experience the highest risks and is estimated to be larger than 200 million RMB. Hebei, Shandong have the most extended infrastructure assets in China. The railway in Shanxi and Sichuan are vulnerable to rainfall-induced hazards, as shown in Fig. 5. When looking at EAD per kilometre for each province, the provinces in Southwest China, such as Sichuan and coastal provinces (e.g. Guangdong, Fujian, and Hainan), have the highest risks. The total EAD and EAD per kilometre are high in Sichuan, Shanxi and Guangdong provinces. From the provincial perspective, these two provinces need to allocate more resources to reduce the risk of rainfall-induced hazards.

Fig. 8 Rainfall-induced hazard risk per province using the average vulnerability curve (a) EAD per province (million RMB) and (b) EAD per km of each province (1000 RMB). Fig. A.1 provides a map of provinces of China.
5. Discussion

This study uses multi-source empirical data to assess the vulnerability and risk to railway infrastructure in China associated with rainfall-induced hazards. For this purpose, the damage news information and a custom damage ratio table are used to fit regional and national vulnerability curves. Previous studies (e.g. Quan Luna et al., 2011; Papathoma-Köhle et al., 2012; Silva and Pereira, 2014; Stephenson and D’Ayala, 2014; Tsubaki et al., 2016; Pregnolato et al., 2017) have tried to use empirical data to fit the fragility or vulnerability curves to hazard intensity and object damage ratio in some regions. In these studies, detailed photos (Papathoma-Köhle et al., 2012; Pregnolato et al., 2017) or hazard model results (Quan Luna et al., 2011) are mostly used to drive the hazard intensity, and adequate documentation of the damage and reconstruction cost can be used to calculate damage ratios. Due to the strict requirement of spatiotemporal damage and hazard intensity information, regional and national vulnerability curves to link hazard characteristics and exposures are rare in many regions. This work tries to overcome the universal problem of the lack of detailed vulnerability data. The fitted vulnerability curves are used as the descriptive damage state in the information on damage and precipitation derived from the news and exited precipitation dataset; these data are more easily collected. Combining the fitted vulnerability curve, precipitation product, and railway infrastructure exposure, the estimated risk of the national railway infrastructure, after calibration with a damage length factor, is approximately 3.91 billion RMB. The overall railway infrastructure risk results are broadly correlated with the yearbook average direct economic damage from 2000 to 2017, which is 3.29 billion RMB. The results reveal that vulnerability...
and risk can be estimated accurately using multi-source empirical data.

Several assumptions and limitations are acknowledged in this study. First, for damage records without local precipitation information, we use the maximum daily precipitation 5 days before damage occurrence (M1-5d) along the damaged segment to present the precipitation intensity. However, there exists deviation for the local damage places along with the damaged segments. In addition, the resolution of the CN05.1 precipitation data is too coarse to accurately capture local extreme precipitation events. We hence use the extracted precipitation information from the news to correct the M1-5d. In a certain way, it would decrease the uncertainty and keep the consistency of the precipitation. Second, due to a lack of different railway market values and detailed information on each railway infrastructure, this work uses the railway market value for 200 km/h railways of double tracks as the value for all types of railway infrastructure. This leads to an overestimation of risk because most conventional railway speeds are lower than 200 km/h, and the relative price has a high probability of being lower than 56 million RMB. Post-disaster reconstruction using higher design standards to improve railways' ability to defend against disasters can reduce the risk for future hazards.

From approximate and common news information to national datasets (e.g. railway damage data), the method used in this work can be a new direction to assess vulnerability and risk by combining multiple sources of empirical data. In addition, the low resolution of the spatiotemporal hazard map smooths the extreme values and cannot capture the hazardous damage. Future research needs to develop a high-resolution spatiotemporal hazard map to prevent this issue.
6. Conclusion

In this study, we use multi-source empirical data to assess the vulnerability and risk to railway infrastructure in China associated with rainfall-induced hazards. Regional- and national-level precipitation vulnerability curves are derived based on news information and a custom damage ratio table. Based on precipitation data, fitted vulnerability curves, the market value of railway infrastructure, and a damage length factor, we assess and calibrate the annual direct damage from 2000 to 2017 caused by rainfall-induced hazards to Chinese railway infrastructure.

Due to the spatial unevenness of protection standards, the regional vulnerability curves of railway infrastructure to rainfall-induced hazards show high spatial inconsistency. Railways in South, Southwest, North, East, and Central China are robust to rainfall-induced hazards since higher protection standards have been used to defend the heaviest rainfall. Railways in Northwest and Northwest China are relatively vulnerable to rainfall-induced hazards. In addition, the regional curves generated in this study can be applied in other works after adjusting the length factor based on the methodology illustrated in sec 3.2.

The national railway infrastructure risk is approximately 3.91 billion RMB, and we find that the estimated annual direct damage of railway infrastructure to rainfall-induced hazards increases due to increasing extreme precipitation and railway exposure. Due to the spatially uneven precipitation intensity, exposure distribution and vulnerability curves, the risk in China shows high spatial differences. The heaviest rainfall and high exposure density lead to a high absolute risk to railway infrastructure in South, East and Southwest China, even though they are robust to rainfall-induced hazards. Provinces such as Sichuan and Guangdong have high
absolute and relative risks. For railway infrastructure risk reduction and sustainable
development of railway transportation in China, more attention and high protection standards
need to be allocated to these high-risk areas. This work provides regional and national
vulnerability and risk information for decision-makers.

Code/Data availability

Supporting data are accessible through the associated reference, and the historical railway
damage data used is in supplement material. The data in this study were analysed with Python
package, and the figures were created with ArcViewTM GIS and Python packages. All codes
used in this work are available upon request.

Author contribution

Kai Liu and Weihua Zhu developed the original idea and designed the analyses. Elco Koks
contributed to the study design. Weihua Zhu and Kai Liu conducted the analysis. Weihua Zhu
wrote the original manuscript, and Kai Liu, Ming Wang, Sadhana Nirandjan and Elco Koks
provided comments and revised the manuscript. All the co-authors contributed to scientific
interpretations of the results.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal
relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Natural Science Foundation of China [grant number
41771538]; the National Key Research and Development Plan [grant number 2017YFC1502901]; and the National Key Research and Development Plan [grant number 2018YFC1508802]. The financial support is highly appreciated.

Reference

Cardoso Pereira, S., Marta-Almeida, M., Carvalho, A. C. and Rocha, A.: Extreme precipitation events under climate change in the Iberian Peninsula, Int. J. Climatol., 40(2), 1255–1278, doi:10.1002/joc.6269, 2020.

Diakakis, M., Boufidis, N., Salanova Grau, J. M., Andreadakis, E. and Stamos, I.: A systematic assessment of the effects of extreme flash floods on transportation infrastructure and circulation: The example of the 2017 Mandra flood, Int. J. Disaster Risk Reduct., 47(February), 101542, doi:10.1016/j.ijdrr.2020.101542, 2020.

Editorial Board of China Railway Yearbook, Ed.: China railway yearbook, China Railway Publishing House, Beijing., 2001.

Editorial Board of China Railway Yearbook, Ed.: China railway yearbook, China Railway Publishing House, Beijing., 2009.

Englardt, J., De Moel, H., Huyck, C. K., De Ruiter, M. C., Aerts, J. C. J. H. and Ward, P. J.: Enhancement of large-scale flood risk assessments using building-material-based vulnerability curves for an object-based approach in urban and rural areas, Nat. Hazards Earth Syst. Sci., 19(8), 1703–1722, doi:10.5194/nhess-19-1703-2019, 2019.

Espinet, X., Rozenberg, J., Ogita, K. S. R. S., Singh Rao, K. and Ogita, S.: Piloting the Use of Network Analysis and Decision-Making under Uncertainty in Transport Operations: Preparation and Appraisal of a Rural Roads Project in Mozambique Under Changing Flood Risk and Other Deep Uncertainties., 2018.

Gerald Ollivier, J. S. and N. Z.: High-Speed Railways in China: A Look at Construction Costs., 2014.

Habermann, N. and Hedel, R.: Damage functions for transport infrastructure, Int. J. Disaster Resil. Built Environ., 9(4–5), 420–434, doi:10.1108/IJDRBE-09-2017-0052, 2018.
Huizinga, J., de Moel, H. and Szewczyk, W.: Global flood depth-damage functions: Methodology and the Database with Guidelines., 2017.

Jia, W. and Xuejie, G.: A gridded daily observation dataset over China region and comparison with the other datasets, Chinese J. Geophys., 56(4), 1102–1111, 2013.

Jongman, B., Kreibich, H., Apel, H., Barredo, J. I., Bates, P. D., Feyen, L., Gericke, A., Neal, J., Aerts, J. C. J. H. and Ward, P. J.: Comparative flood damage model assessment: Towards a European approach, Nat. Hazards Earth Syst. Sci., 12(12), 3733–3752, doi:10.5194/nhess-12-3733-2012, 2012.

Kellermann, P., Schöbel, A., Kundela, G. and Thieken, A. H.: Estimating flood damage to railway infrastructure - The case study of the March River flood in 2006 at the Austrian Northern Railway, Nat. Hazards Earth Syst. Sci., 15(11), 2485–2496, doi:10.5194/nhess-15-2485-2015, 2015.

Kok, M., Huizinga, H., Vrouwenfelder, A. and Berendregt, A.: Damage and casualties caused by flooding, 2004.

Koks, E. E., Rozenberg, J., Zorn, C., Tariverdi, M., Vousdoukas, M., Fraser, S. A., Hall, J. W. and Hallegatte, S.: A global multi-hazard risk analysis of road and railway infrastructure assets, Nat. Commun., 10(1), 1–12, doi:10.1038/s41467-019-10442-3, 2019.

Lamb, R., Garside, P., Pant, R. and Hall, J. W.: A Probabilistic Model of the Economic Risk to Britain’s Railway Network from Bridge Scour During Floods, Risk Anal., doi:10.1111/risa.13370, 2019.

Li, L., Zou, Y., Li, Y., Lin, H., Liu, D. L., Wang, B., Yao, N. and Song, S.: Trends, change points and spatial variability in extreme precipitation events from 1961 to 2017 in China, Hydrol. Res., 51(3), 484–504, doi:10.2166/nh.2020.095, 2020.

Liu, K., Wang, M., Cao, Y., Zhu, W., Wu, J. and Yan, X.: A comprehensive risk analysis of transportation networks affected by rainfall-Induced multihazards, Risk Anal., 38(8), 1618–1633, doi:10.1111/risa.12968, 2018a.

Liu, K., Wang, M., Cao, Y., Zhu, W. and Yang, G.: Susceptibility of existing and planned Chinese railway system subjected to rainfall-induced multi-hazards, Transp. Res. Part A Policy
Liu, K., Wang, M. and Zhou, T.: Increasing costs to Chinese railway infrastructure by extreme precipitation in a warmer world, Transp. Res. Part D Transp. Environ., 93(March), doi:10.1016/j.trd.2021.102797, 2021.

Liu, W., Wu, J., Tang, R., Ye, M. and Yang, J.: Daily precipitation threshold for rainstorm and flood disaster in the mainland of China: An economic loss perspective, Sustain., 12(1), doi:10.3390/SU12010407, 2020.

Nadarajah, S.: The exponentiated Gumbel distribution with climate application, Environmetrics, 17(1), 13–23, 2010.

Papathoma-Köhle, M., Keiler, M., Totschnig, R. and Glade, T.: Improvement of vulnerability curves using data from extreme events: Debris flow event in South Tyrol, Nat. Hazards, 64(3), 2083–2105, doi:10.1007/s11069-012-0105-9, 2012.

Petrova, E.: Natural hazard impacts on transport infrastructure in Russia, Nat. Hazards Earth Syst. Sci., 20(7), 1969–1983, doi:10.5194/nhess-20-1969-2020, 2020.

Pregnolato, M., Ford, A., Wilkinson, S. M. and Dawson, R. J.: The impact of flooding on road transport: A depth-disruption function, Transp. Res. Part D Transp. Environ., 55, 67–81, doi:10.1016/j.trd.2017.06.020, 2017.

Quan Luna, B., Blahut, J., Van Westen, C. J., Sterlacchini, S., Van Asch, T. W. J. and Akbas, S. O.: The application of numerical debris flow modelling for the generation of physical vulnerability curves, Nat. Hazards Earth Syst. Sci., 11(7), 2047–2060, doi:10.5194/nhess-11-2047-2011, 2011.

Sande, van der and C.J: River flood damage assessment using IKONOS imagery, Eur. Comm. Jt. Res. Centre, Nat. Hazards Unit – Floods, Ispra (Va), Italy European Comm. Jt. Res. Centre, Nat. Hazards Unit – Floods, Ispra (Va), Italy, (January 2001), 2001.

Shi, J., Cui, L., Wen, K., Tian, Z., Wei, P. and Zhang, B.: Trends in the consecutive days of temperature and precipitation extremes in China during 1961–2015, Environ. Res., 161(July 2017), 381–391, doi:10.1016/j.envres.2017.11.037, 2018.

Silva, M. and Pereira, S.: Assessment of physical vulnerability and potential losses of buildings
due to shallow slides, Nat. Hazards, 72(2), 1029–1050, doi:10.1007/s11069-014-1052-4, 2014.

Stephenson, V. and D’Ayala, D.: A new approach to flood vulnerability assessment for historic buildings in England, Nat. Hazards Earth Syst. Sci., 14(5), 1035–1048, doi:10.5194/nhess-14-1035-2014, 2014.

Tsubaki, R., David Bricker, J., Ichii, K. and Kawahara, Y.: Development of fragility curves for railway embankment and ballast scour due to overtopping flood flow, Nat. Hazards Earth Syst. Sci., 16(12), 2455–2472, doi:10.5194/nhess-16-2455-2016, 2016.

Ward, P. J., Jongman, B., Weiland, F. S., Bouwman, A., Van Beek, R., Bierkens, M. F. P., Ligtvoet, W. and Winsemius, H. C.: Assessing flood risk at the global scale: Model setup, results, and sensitivity, Environ. Res. Lett., 8(4), doi:10.1088/1748-9326/8/4/044019, 2013.

Yatagai, A., Arakawa, O., Kamiguchi, K., Kawamoto, H., Nodzu, M. I. and Hamada, A.: A 44-year daily gridded precipitation dataset for Asia based on a dense network of rain gauges, Sci. Online Lett. Atmos., 5(1), 137–140, doi:10.2151/sola.2009-035, 2009.

Zhang, X., Su, Z., Lv, J., Liu, W., Ma, M., Peng, J. and Leng, G.: A set of satellite-based near real-time meteorological drought monitoring data over China, Remote Sens., 11(4), 1–12, doi:10.3390/rs11040453, 2019.

Zhao, J., Liu, K. and Wang, M.: Exposure analysis of Chinese railways to multihazards based on datasets from 2000 to 2016, Geomatics, Nat. Hazards Risk, 11(1), 272–287, doi:10.1080/19475705.2020.1714753, 2020.
Appendix

Fig. A.1 Map showing the distribution of Chinese provinces. The China Provincial Map layer comes from the Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences, which is accessible from the Resource and Environment Data Cloud Platform (http://www.resdc.cn/, last access: 19 May 2020).

Table A1 (a) Damage ratio table

| Element          | Unit cost ratio | Damage state | Damage ratio | Description                                                                 |
|------------------|-----------------|--------------|--------------|----------------------------------------------------------------------------|
| Bridges/viaducts | 0.8176          | Severe       | 0.5396-0.8094| Almost components destruction: superstructure, bearing substructure and accessory structure damage |
|                  |                 | Moderate     | 0.2698-0.5396| Two-components destruction: superstructure, bearing substructure and accessory structure damage |
### Table A1 (b) Damage ratio table

| Element            | Unit cost ratio | Damage state | Damage ratio | Description                                         |
|--------------------|-----------------|--------------|--------------|-----------------------------------------------------|
| Tunnels            | 0.8095          | Total        | 0.8095       | Total damage                                        |
|                    |                 | Severe       | 0.5379-0.8069| Almost components destruction: the                   |
|                    |                 | Moderate     | 0.0178-0.357 | Power supply equipment damage                        |
|                    |                 | Light        | 0.0005-0.0178| Catenary pillar destruction                         |
| Component | Damage | Probability | Description |
|-----------|--------|-------------|-------------|
| Track     |        | 0.0794      |            |
|           | Total  | 0.0822      | Total damage |
|           | Severe | 0.0542-0.0814 | Near-failure of components: sleepers, rail, track bed |
|           | Moderate | 0.0271-0.0542 | Two-component failure: sleepers, rail, track bed |
|           | Slight | 0.0008-0.0271 | Single-component failure: sleepers, rail, track bed |
| Signalling and communications | 0.0476 |            | |
|           | Total  | 0.0479      | Total damage |
|           | Severe | 0.0316-0.0475 | Near-destruction of components: digital tuning and TDCS equipment |
|           | Moderate | 0.0158-0.0316 | One-component destruction: digital tuning and TDCS equipment |
|           | Slight | 0.0005-0.0158 | Communication equipment interrupted |
| Electrification | 0.0635 |            | |
|           | Total  | 0.0548      | Total damage |
|           | Severe | 0.0362-0.0542 | Power supply equipment damage and Catenary pillar destruction |
|           | Moderate | 0.0181-0.0362 | Power supply equipment damage |
| Light   | 0.0005-0.0181 | Catenary pillar destruction |
|---------|---------------|-----------------------------|