國立臺灣大學工學院土木工程學系

碩士論文

Department of Civil Engineering
College of Engineering
National Taiwan University

Master Thesis

# 鐵路高風險駕駛行為與路段辨識模組開發

# Development of High-Risk Driving Behavior and Section Identification Modules for Railway System

陳昱甫

Yu-Fu Chen

指導教授：賴勇成 博士

Advisor: Yung-Cheng Lai, Ph.D.

中華民國 109 年 8 月

August 2020

doi:10.6342/NTU202001876

# 國立臺灣大學碩士學位論文
# 口試委員會審定書

鐵路高風險駕駛行為與路段辨識模組開發

Development of High-Risk Driving Behavior and Section Identification Modules for Railway System

本論文係 陳昱甫 君 (R07521506) 在國立臺灣大學土木工程學系完成之碩士學位論文，於民國 109 年 7 月 10 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

| | |
|---|---|
| 賴勇成（指導教授） | 賴勇成 |
| 宋鴻康 | 宋鴻康 |
| 施文雄 | 施文雄 |
| 鍾志成 | 鍾志成 |
| 孫千山 | 孫千山 |
| 系主任 謝尚賢 | 謝尚賢 |

## 致謝

時至今日，我仍難以想像我能夠圓滿完成研究所的課業，用這一篇論文來為 20 年來的學校生活畫下完美的句點。雖然我從小就對鐵路有興趣，但當初的興趣也僅止於對於在兩條平行線上呼嘯而過的龐然大物之憧憬，從來沒有想過要將之作為未來的志業。進入台大土木系後，是賴勇成教授讓我重新燃起對鐵路的熱愛，也要感謝賴老師這兩年來的諄諄教誨，引領我在研究的路途上緩慢但確實地前行，當我迷惘時將我拉出泥淖。除此之外，我也要特別感謝機務處的黃柏文學長、許洋豪前輩，以及行技科的各位先進們，在工作忙碌時仍抽空協助我釐清研究細節、取得資料，並回答我的細碎問題。我也要感謝宋鴻康處長、施文雄副組長、鍾志成主任以及孫千山組長對論文的寶貴意見，還有許聿廷教授、朱致遠教授及陳柏華教授在課程上的教導，讓我能夠在研究的路途上逐漸成長茁壯。

我也必須感謝在研究所一同奮戰的夥伴們。感謝任宏、明華、子皓學長，雖然一同在學校的時間不多，但你們的背影是我不斷追求的目標。感謝敬恆、孟儒學長，你們在努力向前邁步的同時，仍不忘拉我一把。感謝映均學姊，讓我在完成論文的途中與我一同分擔老師的進度壓力。感謝志恆、欣承及功囷，你們幫我承擔了研究以外的各項繁雜事務，希望你們不要把我當作只會把麻煩事丟給學弟的邪惡學長。感謝在 316 室一同奮戰的譽仁學長、明儀學姊、浩雅學姊、冠頡、文心、錦鴻等人，與你們一起的歡笑與淚水是我前進的動力。最後要感謝的是我的家人，感謝你們作為一個避風港一直在我背後支持著我。

感謝其他我無法一一列出的前輩與朋友。最後我只能用一句老套的話來總結我的感謝之意：「要感謝的人太多了，那就謝天吧！」（陳之藩，1962）

陳昱甫　謹誌　於土木工程學系系館　2020.8.12




doi:10.6342/NTU202001876

## 摘要

普悠瑪事故發生後，鐵路安全的議題再次浮上檯面，其中風險評估對鐵路安全至關重要。本研究提出了一個自動化程序，讀取列車自動保護系統（Automatic Train Protection，ATP）之行車記錄以辨識六種高風險駕駛行為。所提出的模組不僅能應用於司機員安全表現評估，亦能針對高風險行為發生之路段進行分析。除此之外，更提出一高風險駕駛行為綜合風險指標，整合六種高風險行為表現以分析並比較不同司機員與路段之間的風險高低程度。在案例分析中，本研究利用實際的行車資料，分別對司機員與路段進行分析，結果顯示若以高風險駕駛行為而言，高達 74%的總體風險來自於 20%的司機員；又以路段而言，高達 80%的總體風險來自於 15%的高風險路段。本研究的模組成功的根據 ATP 行車記錄辨識高風險駕駛以及路段，可便於營運單位對症下藥，有效地提升鐵路系統安全性。



**關鍵字**：鐵路運輸、鐵路安全、自動列車保護、駕駛行為、司機員評估、路段評估、風險指標





doi:10.6342/NTU202001876

## ABSTRACT

Risk assessment is an important process for railway safety. This research proposes an automatic process to access operational record from Automatic Train Protection (ATP) system, and identifies six high-risk driving behaviors. With the integrated risk index for driving behaviors (IRIDB), the proposed modules can successfully identify high-risk drivers and sections in the railway system. Empirical study demonstrates that 20% of high-risk drivers contribute to 74% of the total risk, while 15% of high-risk sections contribute to 80% of the total risk. The proposed modules identify the drivers and sections with high risk to enable the operators of railway systems to take countermeasures, thereby enabling them to efficiently improve the safety of railway systems.



**Keywords**: *Rail Transportation, Railway Safety, Automatic Train Protection, Driving Behavior, Driver Assessment, Section Assessment, Risk Index*





doi:10.6342/NTU202001876

# TABLE OF CONTENT





doi:10.6342/NTU202001876

doi:10.6342/NTU202001876

doi:10.6342/NTU202001876

## LIST OF FIGURES





doi:10.6342/NTU202001876

doi:10.6342/NTU202001876

doi:10.6342/NTU202001876

## LIST OF TABLES





doi:10.6342/NTU202001876

# CHAPTER 1 INTRODUCTION

## 1.1 Background

The overspeed derailment of the Puyuma express train is the deadliest railway accident in Taiwan for the past 20 years. According to the investigation report, switching off of the automatic train protection (ATP) system is one of key contributing factors of the accident (Executive Yuan 1021 RAAIT, 2018).

The Puyuma derailment revealed that no safety equipment can guarantee 100% safety without safety culture and corresponding procedures. By contrast, safety culture and corresponding procedures are the key to the safety of railway systems (Jong et al., 2020). Taiwan Railways Administration (TRA), the operator of conventional railways in Taiwan, has long been criticized for lacking safety cognition in its typical practices. After the Puyuma derailment, the unsafe practices should be improved for safety purposes.

The main purpose of the ATP system is to make sure that trains follow the signals and the speed restrictions. ATP can help drivers identify correct signal aspect and ensure brake are applied properly, if drivers are unconscious due to other factors. The two main functions of ATP are monitoring the train speed and providing the driver information such as the permitted speed, distance to next signal (TRA Electrical Engineering Department, 2005). Although the ATP system is introduced to reduce railway accidents, it may, reversely, increase the risk of train operations.

Two types of human factors, distrust of ATP and over-trust of ATP, that are potential hazards to railway system safety may arise because of the complete function of ATP. Both of them increase risk of train operations (ITSRR, 2005).

The assessment of drivers' performance is one of the key approaches to ensure safety operations. Figure 1-1 shows that on October 14, 2018, one week before the



doi:10.6342/NTU202001876

accident, the driver drove the same train through the accident site (near Xinma Station). The driver did not start to decelerate the train before hearing the warning from the ATP system. By contrast, a normal driver would decelerate the train at Wulaokeng Bridge (as shown in Figure 1-2). On October 21, 2018, when the accident driver switched off the ATP system, he did not decelerate the train at all, thereby resulting in the overspeed derailment (Executive Yuan 1021 RAAIT, 2018). A suspicious aspect is that drivers tend to operate trains by relying on ATP, which is supposed to be a protection system instead of a guiding system. TRA drivers should operate trains by following wayside signals. However, some drivers have been found to rely on the continuous and convenient information on speed limit from the display panel of the ATP system. Consequently, the current research proposes an automatic process to access ATP records to identify high-risk driving behaviors.

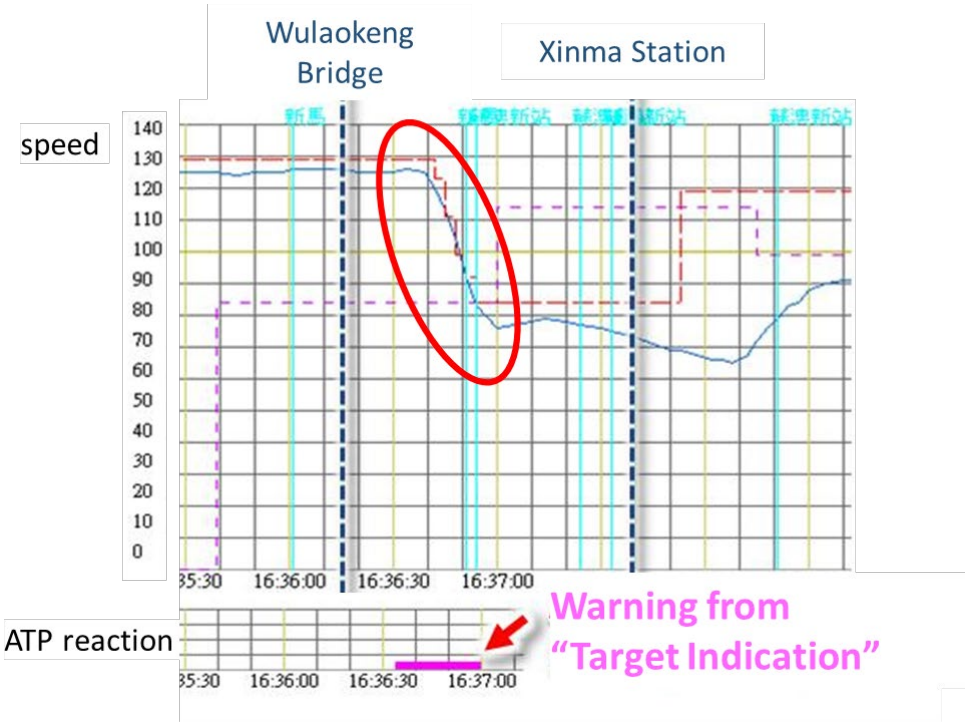


Figure 1-1: Record of the accident driver on October 14, 2018

(one week before the derailment)



doi:10.6342/NTU202001876

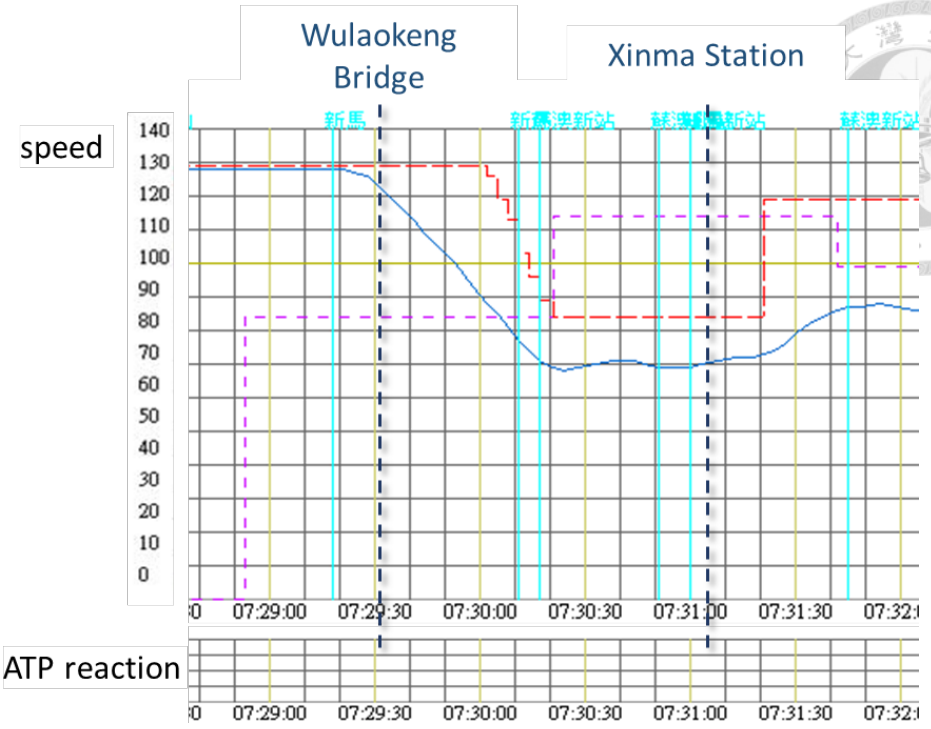


Figure 1-2: Record of a normal driver

The Puyuma derailment revealed two safety concerns existing in TRA: hardware problems and human factors. The solution for the former is to improve the hardware of the ATP system, which is out of scope of this research. This research focuses on the solution to the latter concern: to improve the driver performance in train operations. This objective can be achieved by evaluating the driver performance and retraining those with worse performance.

In the past, the assessment of driver performance was a manual process. Driving instructors who were responsible for assessment checked the figures output from ATP one by one and inspected if any driver had violated regulations. This process is not only inefficient but also likely to overlook some small mistakes. Furthermore, it's impossible for several experienced drivers to monitor about 1,000 trains operated every day. Since early 2020, TRA introduced an overspeed detection system for driver assessment. The system can process ATP data and identify approach signal overspeed. However, the overspeed detection system is only designed to identify one specific behavior.



doi:10.6342/NTU202001876

The other focus of safety improvement is the high-risk sections. After the Puyuma derailment, TRA also planned to reconstruct sections with sharp curves in the network of TRA. However, there is no practical methodology to identify sections with high risks.

As mentioned in previous paragraphs, it is urgent to develop modules automatically identifying high-risk driving behaviors and sections, so that the operators can take countermeasures to improve the safety of train operations.

## 1.2 Research Objectives

As was mentioned in the previous section, this research aims to develop a process to identify the high-risk driving behaviors based on the ATP driving record data, and calculates the frequency of the high-risk driving behaviors for each driver and section. An integrated index of high-risk driving behaviors is also proposed to evaluate the performance of the drivers and the safety of the sections. At first, ATP record figures are input and transferred into numerical data for further analysis. Every ATP driving record data is categorized by drivers and sections separately afterwards. With the pre-processed ATP driving record data, six identification modules are established to discover six different types of high-risk driving behaviors. Moreover, this research also establishes an index that mixes all six high-risk driving behaviors for safety assessment. One case study involving 540 drivers and the other involving 142 sections in TRA are applied by this assessment process to evaluate safety on practical scenarios. The evaluation of risk index reveals the drivers and sections with higher risk, and TRA can take actions to improve safety accordingly.

## 1.3 Contribution Summary

The contributions of this research could be summarized into the following three points:

(1) Define and discover six high-risk driving behaviors that may be latent risk to future



doi:10.6342/NTU202001876

accidents.

Although no injuries or property damages are caused by these high-risk driving behaviors, they probably trigger to serious accidents in combination with other failures. Once these high-risk driving behaviors are identified and improved, possibility of accidents would also decrease.

(2) Develop an automatic process to identify high-risk driving behaviors with the ATP record figures.

The framework of high-risk driving behavior identification process, including the figure transformation and the analysis modules for the drivers and sections, are fully automatic processes. No extra human resource is required in this process. It not only greatly improves the manual assessment process in TRA but also provides more functions for risk evaluation.

(3) Propose an integrated risk index for quantifying high-risk driving behaviors for each driver and section.

An integrated risk index of high-risk driving behaviors and sections can provide a quantified score of the risk to high-risk driving behaviors and sections. It can reflect danger and urgency parts of the whole railway system. With this index, operators can establish an improvement plan accordingly.

## 1.4 Thesis Organization

This thesis is organized into five chapters. CHAPTER 1 describes the background and objectives of this research. Afterwards, literature related to the railway driver performance and safety and evaluation of risk index is reviewed in CHAPTER 2. CHAPTER 3 introduces the framework and modules of the high-risk driving behaviors and high-risk sections identification and risk evaluation process. Two case studies, one for analysis of drivers; the other for analysis of sections, are demonstrated in



doi:10.6342/NTU202001876

CHAPTER 4. Finally, CHAPTER 5 concludes this research and shows the possible ways of future work. Figure 1-3 summarizes the organization of this research.

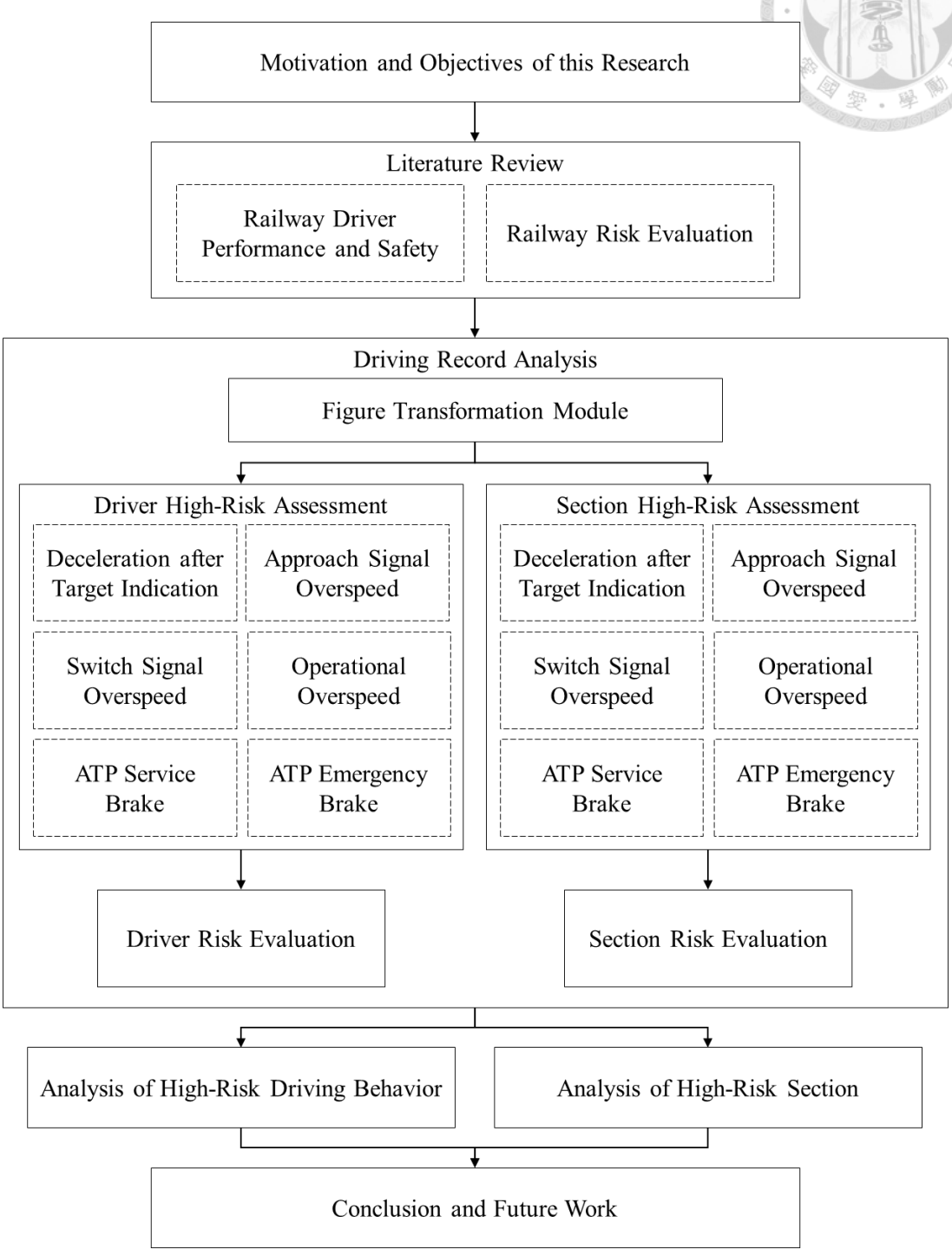


Figure 1-3: Thesis Organization



doi:10.6342/NTU202001876

# CHAPTER 2 LITERATURE REVIEW

The purpose of this chapter is to review literature on driver performance and risk evaluation. It begins with reviewing researches about analysis on railway driver performance and safety. In Section 2.1, three types of methodologies about railway driver performance: human factor analysis, accident review, and driving data analysis are reviewed in Sections 2.1.1 to 2.1.3 separately. Afterwards, section 2.2 reviews the literature about railway risk evaluation and risk index. Finally, in Section 2.3, the reviewed literature is summarized, and innovation of this research is also discussed in the section.

## 2.1 Safety and Performance of Railway Drivers

Numerous studies have discussed the railway safety. This section focuses on the studies related to safety and performance of railway drivers. These studies can generally be categorized into three types according to research methodology: human factor analysis, accident data analysis, and train operation data analysis, discussed in Sections 2.1.1 to 2.1.3 respectively.

### 2.1.1 Human Factor Analysis

In terms of railway driver safety and performance, the methodologies of human factor analysis are mostly used for research (Wilson and Norris, 2005). Typically, these studies use experiments or interviews to evaluate driving performances and one or some factors that may affect driving safety. Thereafter, the relationship and effect between them are analyzed. Among them, workload and fatigue are popular factors that are of interest to researchers. Dorrian et al. (2007) used driving simulators to compare driving behaviors of nighttime shift after two consecutive nighttime shift to daytime shift after normal sleep. Jay et al. (2008) observed driver fatigue of the driver relays (alternative shifts of two set of crews) by assessing fatigue levels before and after each shift. Dorrian



doi:10.6342/NTU202001876

et al. (2011) assessed fatigue levels of drivers, controllers, and maintenance crew before and after the shifts. Apart from workload and fatigue, driver distraction and inattention are other key factors worth studying. Naweed (2013) collected several signals passed at danger (SPAD) scenarios and analyzed SPADs with psychology models to find the relationship between SPAD-risk and distraction. There are also studies focus on drivers' personality. Hickey and Collins (2017) examined driving performance and driving-related incidents. Other studies have integrated several of the preceding factors to establish a considerably general model for driving behavior. Myrtek et al. (1994) assessed physical and mental conditions of train drivers of different type of railway systems, and figured out factors that may affect driving performance. McLeod et al. (2005) constructed a psychological situational model for railway drivers to describe perception, decision and action of the drivers. Naweed (2014) also used psychological methods to investigate driver characteristics, parameterize conventional driver behavior, and developed a model of train driving behavior. Tabai et al. (2018) investigated drivers' cognitive and demographic characteristics and their effect on accidents.

### 2.1.2 Accident Review

Another type of method to assess driver safety and performance is based on past accidents. These studies have often been based on accident datasets and reports to analyze human factors that contributed to accidents. Reinach and Viale (2006) used Human Factors Analysis and Classification System (HFACS) to establish a complete and formal structure of accident investigation and analysis, and analyzed some remote control locomotive accidents and incidents according to the proposed structure. Baysari et al. (2008) also applied HFACS to 40 accident investigation report, discussing reasons to human errors and approaches for prevention. Kyriakidis et al. (2015) integrated previous literature and hundreds of accidents or incidents worldwide, induced



doi:10.6342/NTU202001876

contribution factors to accidents, and analyzed effect and relationship of the accidents. An index was also proposed to evaluate behavior and performance of drivers. Guo et al. (2016) surveyed personality of the train drivers and analyzed the relationship between personality and accident involvement, sobriety test and closed-circuit television records using correlation analysis and Poisson regression analysis. Punzet et al. (2018) analyzed the SPAD reports and researched in the human factors of SPADs. Besides researching with conventional accident analysis models, more and more researches turned to apply some novel machine learning methods to find latent relationships that were ignored previously. Mirabadi and Sharifian (2010) applied data mining methodologies to accident dataset and revealed relationships between accidents and various factors. Figueres-Esteban et al. (2016) used artificial intelligence to analyze text-based records and made the records applicable for railway safety analysis. Kaeeni et al. (2018) integrated three kinds of machine learning skills: artificial neural network, Naïve Bayes, and decision tree to construct a risk prediction model of derailment accidents. Rashidy et al. (2018) collected many factors seemingly not relevant to SPAD and examined the factors by big data analysis. A dataset that can automatically upload and pre-process data collected is also established.

### 2.1.3 Driving Data Analysis

With the development of big data analysis skills, some studies have started to apply train driving records to assess driving performance and safety. Sun (2010) applied ATP record data, counting two types of behavior defined as “near miss.” Statistical analysis was also performed to analyze the relationship between the near miss behaviors and some factors including personalities, train types, track conditions, and environments. Zhao et al. (2018) used data from conventional signal systems to detect red signal approach for future studies related to SPAD. El Rashidy et al. (2018) applied driving



doi:10.6342/NTU202001876

data from on-train monitoring recorders to establish an index to evaluate train driving performance.

## 2.2 Railway Risk Evaluation

A variety of techniques can be used to evaluate the risk in railway systems, and each of them has different characteristics and suitable conditions. These techniques can be categorized into three types: quantitative techniques, semi-quantitative techniques, and qualitative techniques (ISO, 2010).

One of the most well-known concepts of quantitative risk analysis is shown as equation (1) (Boehm, 1989):

$$Risk = Frequency \times Consequence \tag{1}$$

Though the concept was firstly used in software risk assessment, it is gradually applied in a variety of fields, including railway operation. Macciotta et al. (2016) used Monte Carlo simulation to estimate the risk of a slope section. Leitner (2017) established a model for railway risk assessment based on the accident database. In the model, probability and consequence are calculated by the fault tree analysis and event tree analysis, respectively.

Although the quantitative risk analysis can calculate the precise value of risk, the calculation is complicated and arbitrary. Moreover, strong simplifications and assumptions are usually necessary for quantitation, thus some important factors may be neglected. Instead, qualitative or semi-quantitative risk analysis can be more comprehensive, wherein more latent factors that are influential to risk can be considered. Furthermore, qualitative or semi-quantitative risk analysis required less resource so that it is more practical to adapt with (Aven, 2008).

Practically, the risk matrix is often used in risk assessment for railway systems, and it is a methodology applicable for both qualitative and semi-quantitative risk



doi:10.6342/NTU202001876

analysis. In EN 50126 standard, risk matrix is referred as a technique to evaluate the risk levels (CENELEC, 1999). Numerous railway operators have adopted risk matrices as their risk assessment method, such as London Underground, Hong Kong MTR, Taiwan High Speed Rail, Taipei Metro (NTURTRC, 2012).

Although the risk matrix is a mature technique for risk assessment in railway, the analysis is on the basis of apparent deaths and injuries. The hazard events with less risk is not considered in the risk matrix. As defined by Health and Safety Executive (2004), adverse events can be divided into accidents and incidents. Accidents represent events where deaths, injuries or ill health are caused, while incidents include "near miss" and "undesired circumstance," which are events or circumstances potentially result in injuries or ill health but no actual harm is caused. However, this research focuses on "high-risk driving behaviors of drivers," and such events are leading indicators before accidents and incidents happen. Therefore, common methods for risk evaluation of transportation is not applicable in this research.

In contrast with the risk evaluation methods, little literature proposed methods for near miss risk evaluation. Kaplan (1990) proposed a method based on the concept of event tree analysis and Bayesian theorem to evaluate the risk of near miss events. Gnoni and Lettera (2012) proposed a model to calculate a risk index for near miss events with consideration of the hazard of near miss events, effects of location, and feasibility of countermeasures. McKinnon (2012) considered the risk as a combination of probability, frequency, and severity, where probability is the likelihood that near miss events become accidents, frequency is the number of near miss events in a time, and severity is the consequence of the worst cases. Conclusively the risk index can be calculated by multiplication of these three elements, as shown in equation (2).

$$Risk = Probability \times Frequency \times Severity \quad (2)$$



doi:10.6342/NTU202001876

Although some methodologies for near miss risk assessment have been proposed, all of them are inapplicable due to the limitation of accessible data. Consequently, a new risk evaluation method for near miss events is necessary in this research.

MacKenzie (2014) suggested some methods for establishing a risk index in different areas: the moments of distribution, the quantiles of a distribution, the disutility functions, and the functions of indicators or factors. For the first two methods, there is no apparent theory to determine which moments or quantiles should be chosen. For the third method, there is no sufficient data for calculating the disutility function. Moreover, the research aims to propose an integrated risk index of several high-risk driving behaviors. Hence, the last method, the functions of indicators or factors, is the most appropriate method for constructing a risk index in this research.

## 2.3 Summary of Literature Review

In this chapter, literature with regard to the driver performance and safety is reviewed. In summary, researches of human factor analysis assessed the driver behavior via ergonomic and psychological methods. The disadvantages of these studies include a high threshold of technique. Professional psychologists are necessary for the majority of these studies. Moreover, the majority of these studies are interview-based, thereby making it difficult to quantify objectively. For accident review researches, accident database is collected to analyze relationship between accidents and driver performance. The major problem is that accidents are lagging indicators, analyses can only be done after damages have happened instead of preventing damages from happening. Studies of the third methodology, driving data analysis, collect historical driving data and identify abnormal driving actions for further analysis. This methodology emerged with the development of big data analysis. Driving data analysis can directly observe driver behavior and analyze it through quantification analysis. Consequently, driving data



doi:10.6342/NTU202001876

analysis is chosen as the methodology in the research.

By reviewing the literature about driving data analysis, it is found that few previous studies have used driving data analysis to assess driver performance and safety. Some of these researches identified specific actions manually, while the other researches did not discuss the relationship between signal information, speed limit and driver performance. This research applies the driving record from ATP, which is equivalent to the European Train Control System (ETCS) Level 1, to identify the high-risk driving behaviors. This research proposes an automatic process to analyze the interaction between driver and signal information. Moreover, this research not only focuses on the high-risk driving behaviors of each driver, but also focuses on the sections where these high-risk driving behaviors happens.

Furthermore, this research also proposes a framework of risk index evaluation for the high-risk driving behaviors. Previous studies and standards for railway risk evaluation only focused on the risk of accidents, and these methods are difficult to be applied for near miss events such as high-risk driving behaviors in this research. This research also proposes a framework of calculating an integrated risk index for events with relative low damages.



doi:10.6342/NTU202001876

# CHAPTER 3 METHODOLOGY

This chapter describes and discusses the methods used in this research. Firstly, Section 3.1 introduces the functions and the control logics of the ATP system in TRA. In the second part of the section, the defects and the safety concerns of ATP are discussed. Sections 3.2 to 3.6 then describe the process of the research in greater detail. Section 3.2 provides a brief overview of the overall process and introduces the modules used in this research, as well as inputs and outputs of each module. Sections 3.3 to 3.6 are detailed description of the modules.

## 3.1 Introduction of ATP System in TRA

Since this research applies the driving record data from ATP, this section begins with introducing the functions of ATP. Then, the safety concerns of ATP, which is the main reason to propose this research, is discussed in the second part of this section.

### 3.1.1 ATP System of TRA

The main function of the ATP system for TRA is to ensure that the train would follow wayside signal and speed limit by monitoring train speed continuously. The equipment of the ATP system consists of two parts: wayside and onboard equipment. Wayside equipment comprises the existing signaling system, encoder and balise (also called transponder). Information on train movement, such as moving authority and speed limit, are encoded by an encoder and sent to the onboard equipment from balise. After receiving information from balise, the onboard computer calculates the permitted speed profile and monitors the train speed according to the received information and pre-input train characteristics.

The function of the ATP system in TRA is equivalent to the ETCS Level 1, which is an intermittently updated continuously supervised train protection system. Although ATP can monitor train speed throughout the journey, signal information can be updated



doi:10.6342/NTU202001876

only on locations where balises are installed. In general, 2 to 3 balises are installed before each signal. When trains pass through a balise, the onboard equipment receives the corresponding signal aspect and rail alignment information of the following block, such as speed limit, information on curves, slopes, turnouts, and distance to the next signal. The received speed limit is the integrated speed restriction that considers curves, slopes, bridges, turnouts, and stations, among others.

Note that unlike a conventional three-aspect speed restriction, the speed restriction of ATP is based on speed profile, which is calculated using the stop point and brake performance of the train (see Figure 3-1). ATP calculates five different speed profiles: target indication, permitted speed, service warning, service brake, and emergency brake. Figure 3-2 and Figure 3-3 illustrates the relationship of the five speed profiles while confronting a stop signal or speed limit reduction. Each of the profiles represents a different level of ATP interference, which is discussed in the following section. ATP uses an odometer to locate the train and monitor train speed corresponding to the current location.

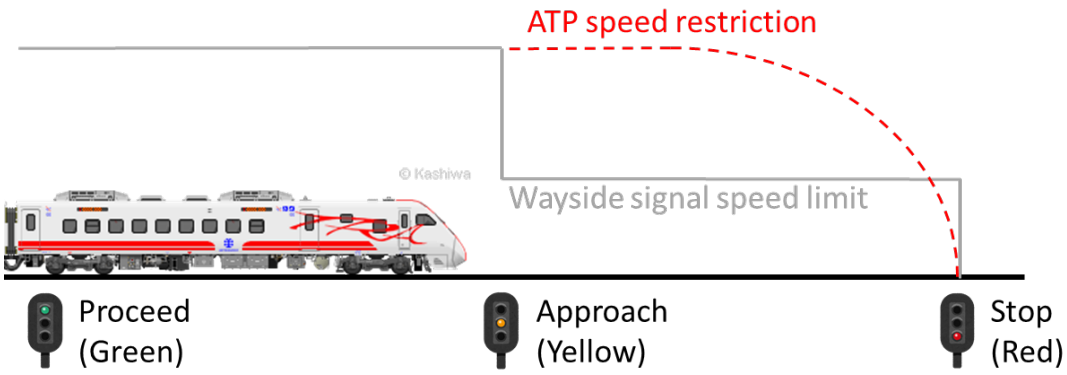


Figure 3-1: Difference of speed restriction between signal and ATP



doi:10.6342/NTU202001876

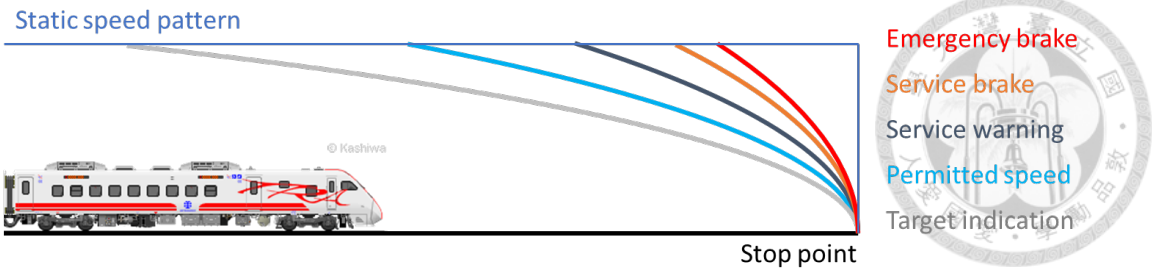


Figure 3-2: Schematic of speed profiles while confronting a stop point

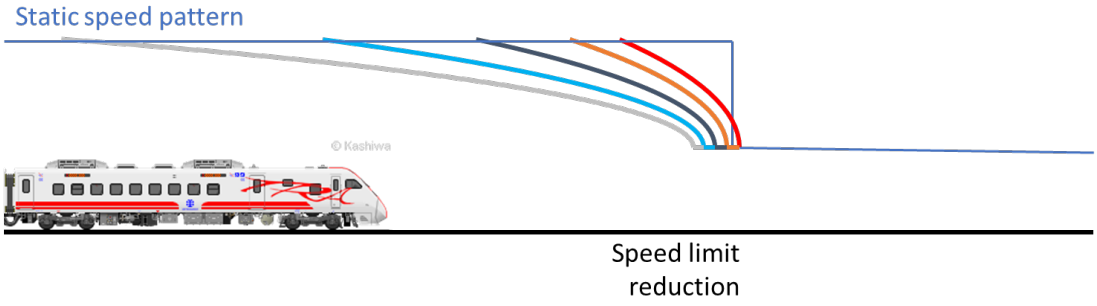


Figure 3-3: Schematic of speed profiles while confronting a speed limit reduction

Figure 3-2 and Figure 3-3 shows the schematic of the five speed profiles calculated by ATP when it detects a stop point or speed limit decrease in front of the train. These figures show that static speed profile is the speed limit received from the wayside equipment, representing the speed limit of the section. “Target indication” is a reminder before the train speed reaches the permitted speed profile. When the train speed reaches the target indication profile, the “train would overspeed in 4 seconds with current speed.” At this time, the ATP monitor will display the target speed, and the ATP buzzer will beep to remind the driver. After 4 seconds with the same speed, the “permitted speed” profile will be reached. Permitted speed profile is the speed limit calculated by ATP, but nothing happens if the train speed exceeds the permitted speed. After an additional second with the same speed, the “service warning” profile is reached. Service warning profile aims to warn the driver that the brake should be applied to prevent ATP brake interference. The ATP buzzer will continue sounding if the train speed exceeds the service warning profile. If the train is operated with the same speed for an additional 9



doi:10.6342/NTU202001876

seconds, then it will reach “service brake” profile. At this time, the maximum service brake is automatically activated by ATP to slow down the train, until the train speed is lower than the permitted speed. If ATP considers that the maximum service brake is not sufficient to stop the train before the stop point, then the “emergency brake” will be applied by ATP. Not until the train stops can the driver release the emergency brake (Bombardier Transportation, 2005).

Although ATP records train operation data continuously and saves all data in a database, all data is encrypted so that the original data is unavailable without keys. The accessible data in this research are figures of driving records. Figure 3-4 demonstrates part of the driving record figure and the information contained in the figure.

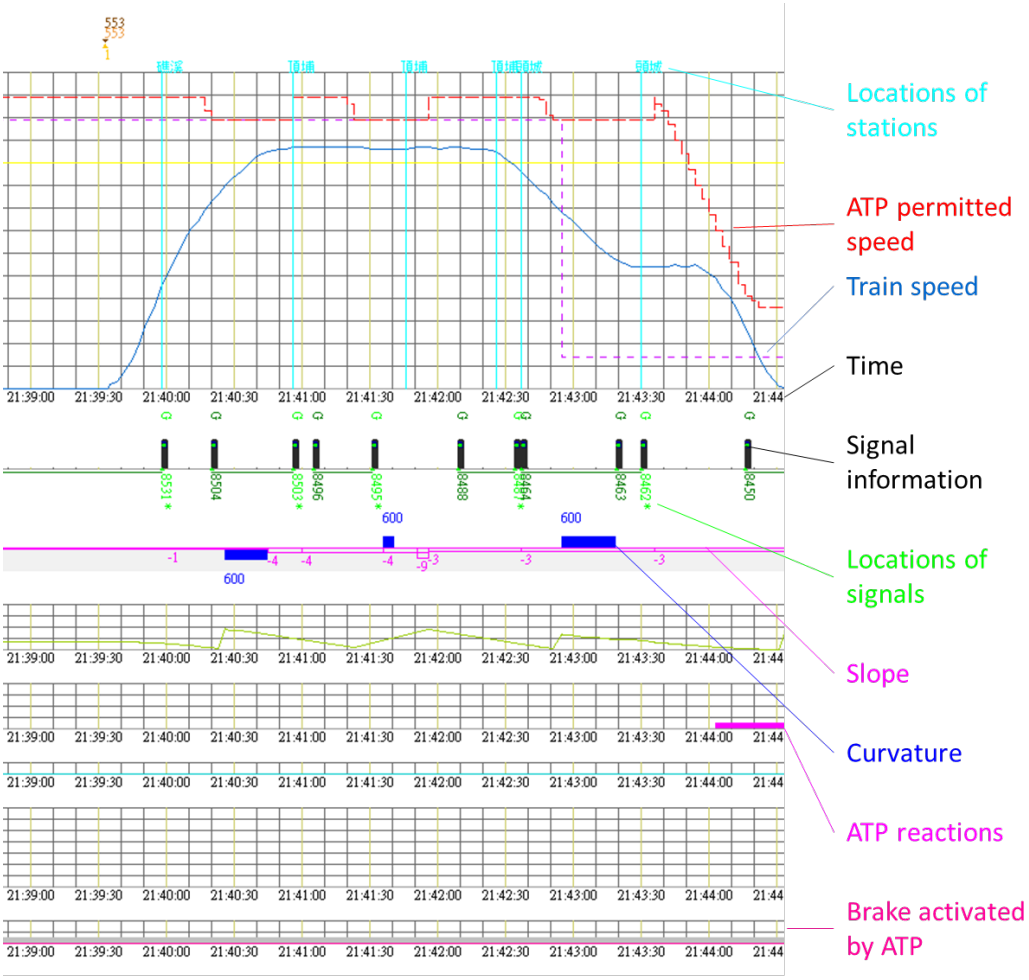


Figure 3-4: Part of the driving record figure and the contained information



doi:10.6342/NTU202001876

### 3.1.2 Safety Concerns of ATP System in TRA

Although ATP provides a complete safety protection of train operation, ATP continues to have some safety concerns. Human factors are key factors that potentially threatens the safety of railway operations. Although ATP has been developed and applied in operation for decades, only a few studies have discussed the effects of ATP on driving performance. The human factors of ATP could be categorized into two types: distrust and over-trust of ATP. Distrust of ATP could result because of two reasons: low reliability of ATP and different control logic between driver and ATP. In 2018, 332 failures were recorded out of 6,530 sets of wayside equipment, and 224 of the failures were the result of unknown reasons (Executive Yuan, 2019). Moreover, when ATP fails, brake is often activated to ensure safety owing to the fail-safe characteristic of ATP. Thus, drivers are often aware of ATP malfunction because of its interference of train operation, thereby deepening the impression that ATP fails frequently. Consequently, drivers become accustomed to the frequent malfunction of ATP. Once problems related to power loss or brake activation occur, drivers tend to blame them on ATP malfunction. By contrast, over-trust of ATP is also a problem, particularly given that ATP is not a cure-all for train operations. ATP cannot handle some situations, such as trespassers and foreign objects on the track. If drivers consider that ATP can prevent any disaster, then they may lose their concentration. Another concern is the effect over driving skill. Given that ATP can integrate every necessary information and generate speed restrictions by considering all conditions, drivers may get used to driving dependent on the information from ATP, thereby gradually losing their ability to judge train operations. Hence, they may be unable to make correct decisions when ATP does not work (ITSRR, 2005).

On the basis of the regulation of ATP usage and management in TRA, ATP cannot



doi:10.6342/NTU202001876

be turned off unless the ATP equipment malfunctions. Moreover, once ATP is turned off, the driver must contact dispatchers and report some necessary information. However, the process after ATP is turned off is not applied properly, thereby resulting in disasters, such as the Puyuma derailment (Executive Yuan 1021 RAAIT, 2018). In addition, the ATP system in TRA is treated as a safety protection system. Drivers should follow the wayside signal instead of ATP onboard information, even though the location of the ATP display panel is merely in the middle of the driving dashboard (TRA, 2018). The complete information and convenience may trigger drivers' reliance on ATP.

### 3.2 Procedure of High-risk driving behavior Analysis

Due to the difficulty of inspection and analysis of ATP data, this research proposes an automatic process to transform the figure into numerical data and further figures out some high-risk driving behaviors. Finally, an integrated risk index for driving behaviors (IRIDB) is evaluated. The process is performed with Python programming language on Spider integrated development environment (IDE). The flowchart of the research method is illustrated in Figure 3-5.



doi:10.6342/NTU202001876

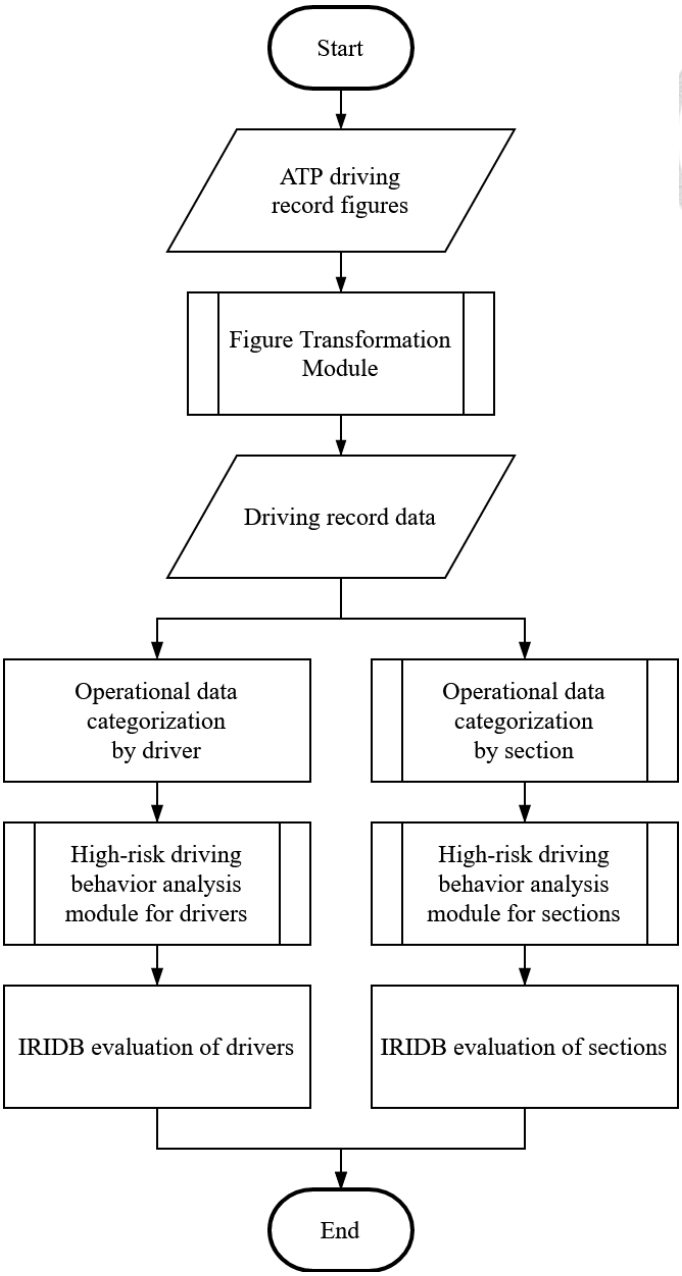


Figure 3-5: Flowchart of the research method



doi:10.6342/NTU202001876

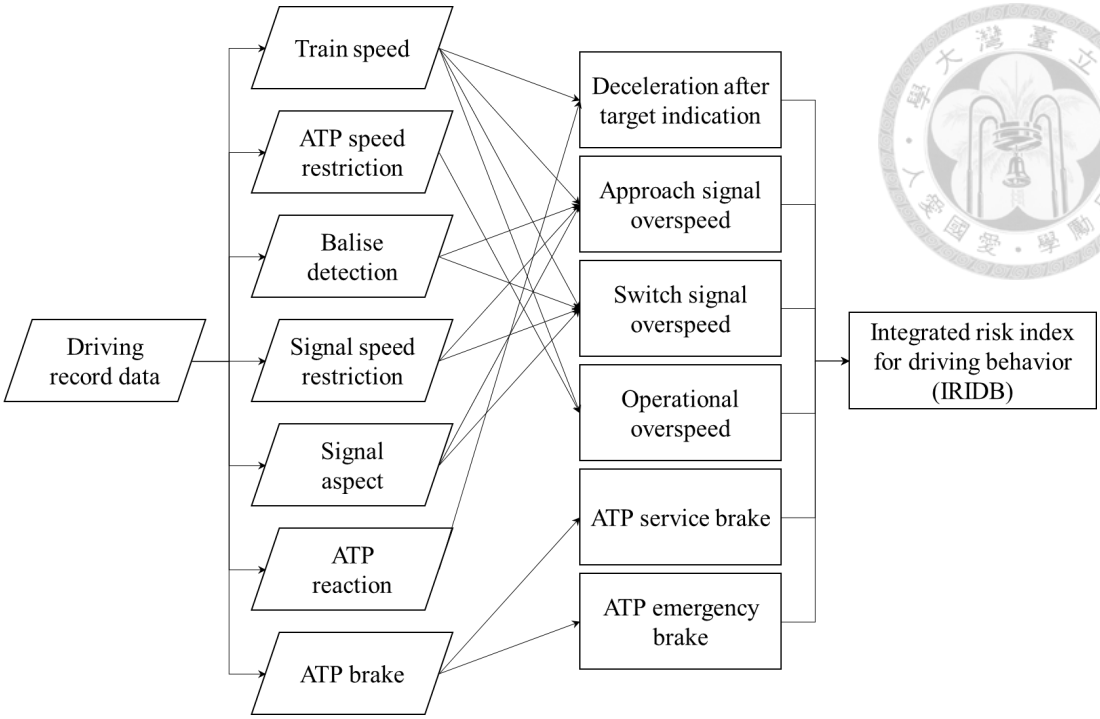


Figure 3-6: The data used for the high-risk driving behavior identification and the risk index evaluation

Six types of high-risk driving behaviors are identified and discussed in this research, namely, “deceleration after target indication,” “approach signal overspeed,” “switch signal overspeed,” “operational overspeed,” “ATP service brake,” and “ATP emergency brake.” Each of them is applicable under different conditions and identified by specific module. Figure 3-6 presents the six high-risk driving behaviors and the data used for identification. Deceleration after target indication is an interaction between ATP and driver, so the train speed and the ATP reaction are applied for identification. With the balise detection, the signal speed restriction and the signal aspect along the whole journey can be inferred. Thus, approach signal overspeed and switch signal overspeed can be identified with combination of the train speed and the signal condition. Lastly, ATP service brake and ATP emergency brake can be identified once the service brake or the emergency brake is activated by ATP. Indicators of the six high-risk driving behaviors for driver and section analysis are listed in Table 3-1. More detailed



doi:10.6342/NTU202001876

information about the high-risk driving behaviors and the process for identification will be introduced in Sections 3.5.1 to 3.5.6, respectively.

Table 3-1: Indicators of high-risk driving behaviors

| High-risk driving behaviors | Indicators | Units |
|---|---|---|
| Deceleration after target indication | Number of behaviors per unit train-distance | times/train-km |
| Approach signal overspeed | Time of behaviors per unit train-distance | seconds/train-km |
| | Time of behaviors × excessive speed per unit train-distance | $\frac{s \cdot km/h}{train\text{-}km}$ |
| Switch signal overspeed | Number of behaviors per unit train-distance | times/train-km |
| | Sum of excessive speed per unit train-distance | $\frac{km/h}{train\text{-}km}$ |
| Operational overspeed | Time of behaviors per unit train-distance | seconds/train-km |
| | Time of behaviors × excessive speed per unit train-distance | $\frac{s \cdot km/h}{train\text{-}km}$ |
| ATP service brake | Number of behaviors per unit train-distance | times/train-km |
| ATP emergency brake | Number of behaviors per unit train-distance | times/train-km |

## 3.3 Figure Transformation Module

The first step of analysis is to read the ATP driving record figures and transform the figures into a numerical array. In this step, imageio, a Python library, is used to read images decrypted by the ATP system and to transform the figures into three-dimensional arrays. The x-axis and the y-axis of the array are identical to the size of the input image in pixels. The z-axis of the array demonstrates the RGB color corresponding to the image. Based on the RGB array constructed, operation data can be obtained by analyzing the relative location of each point in the array. Table 3-2 shows the data



doi:10.6342/NTU202001876

obtained via this process. After obtaining theses data, operating distance of each data point is calculated by integrating train speed over driving time.

Table 3-2: Data obtained via figure transformation module

| Data type | Format |
|---|---|
| date | numerical (yyyymmdd) |
| train number | numerical |
| driver ID | numerical |
| time of operation | numerical (seconds since beginning) |
| train speed | numerical (km/h) |
| ATP speed restriction | numerical (km/h) |
| balise/signal detection | categorical (none, balise, signal) |
| signal speed restriction | categorical (140, 60, 45, 35, 25) |
| type of signal | categorical (none, mainline signal, turnout signal) |
| ATP reaction | categorical (none, target indication, service warning) |
| ATP brake | categorical (none, service brake, emergency brake) |

The 11 types of data listed in Table 3-2 is all information that can be applied for high-risk driving behavior identification. Among the data, the date, the train number, and the driver ID are accessed while inputting ATP record figures, and the time of operation is obtained by counting the pixels from the beginning of the journey. The parts from which the data of Table 3-2 is identified are depicted in Figure 3-7. The top part of the driving record data demonstrates the speed profiles of the train. The blue solid line represents the train speed, while the red dashed line represents the ATP permitted speed profile. In the part below, a sequence of signal-liked patterns show the signal and balise information. The signal-liked patterns represent the locations of the balises, and the signal color reflects signal aspects received from the balise. The “*” mark under the signal-liked patterns indicates positions of the signal machines. The orange part and the blue part in the bottom of the figure show the reaction of ATP. In



doi:10.6342/NTU202001876

the orange part, purple blocks appear during the target indication or the service warning. Once maximum service brake is activated by ATP, a blue solid line will appear. On the other hand, once emergency brake is activated by ATP, a pink solid line will appear in the blue part.

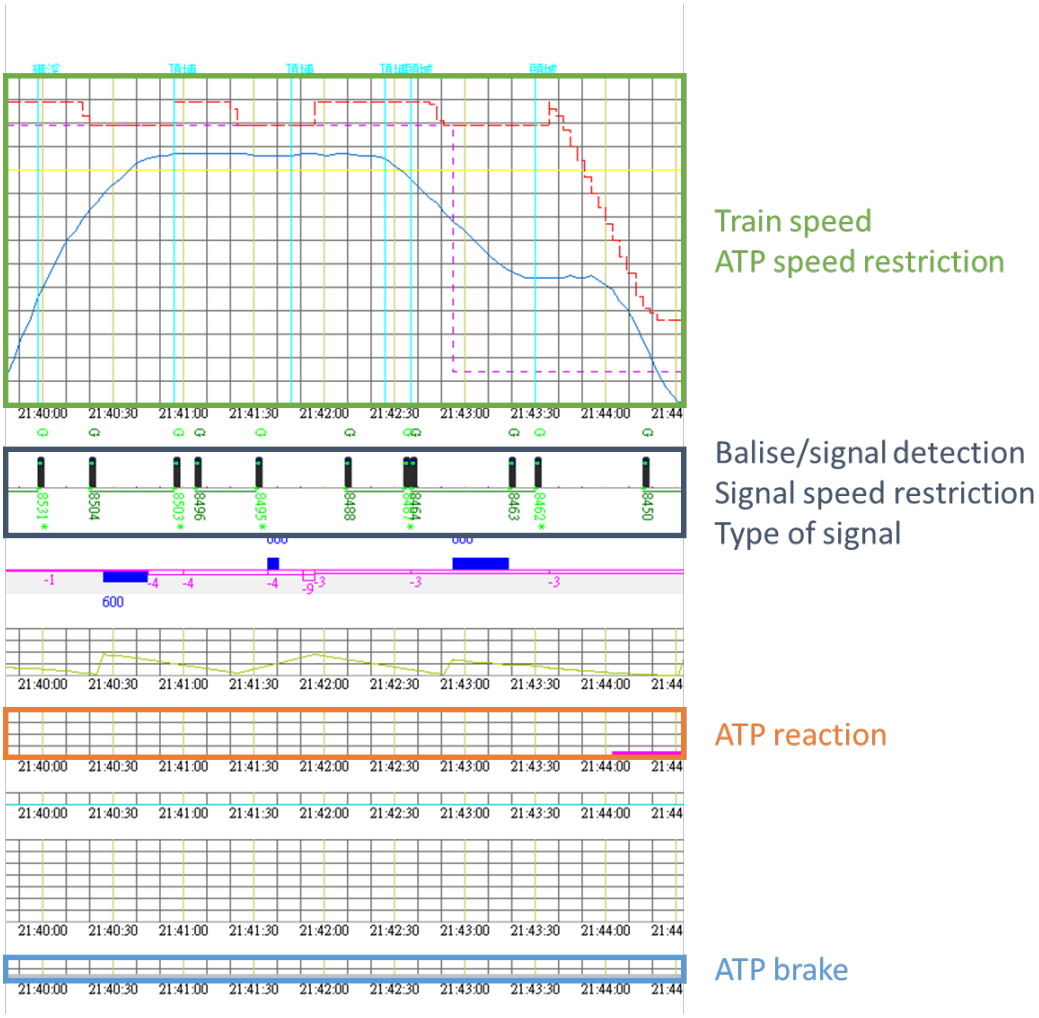


Figure 3-7: Parts used for data identification in this module

With the figure transformation module, the ATP driving record figures are transformed into numerical driving record data, which can be further analyzed in the following sections.

### 3.4 Driving Record Data Categorization

This research not only targets at the drivers who commits high-risk driving behaviors frequently, but also focus on the sections where drivers commit high-risk



doi:10.6342/NTU202001876

driving behaviors frequently. Thus, data categorization by both drivers and sections are necessary for further analysis.

Since the driver ID is accessible from previous modules, the driving record data categorization by drivers is a relatively easy, where only reading the driver ID and categorizing all data accordingly. On the other hand, the process for driving record data categorization by sections is much more complex. Figure 3-8 summarizes the process for driving record data categorization by sections. The thorough description of this process is introduced as follows:



doi:10.6342/NTU202001876

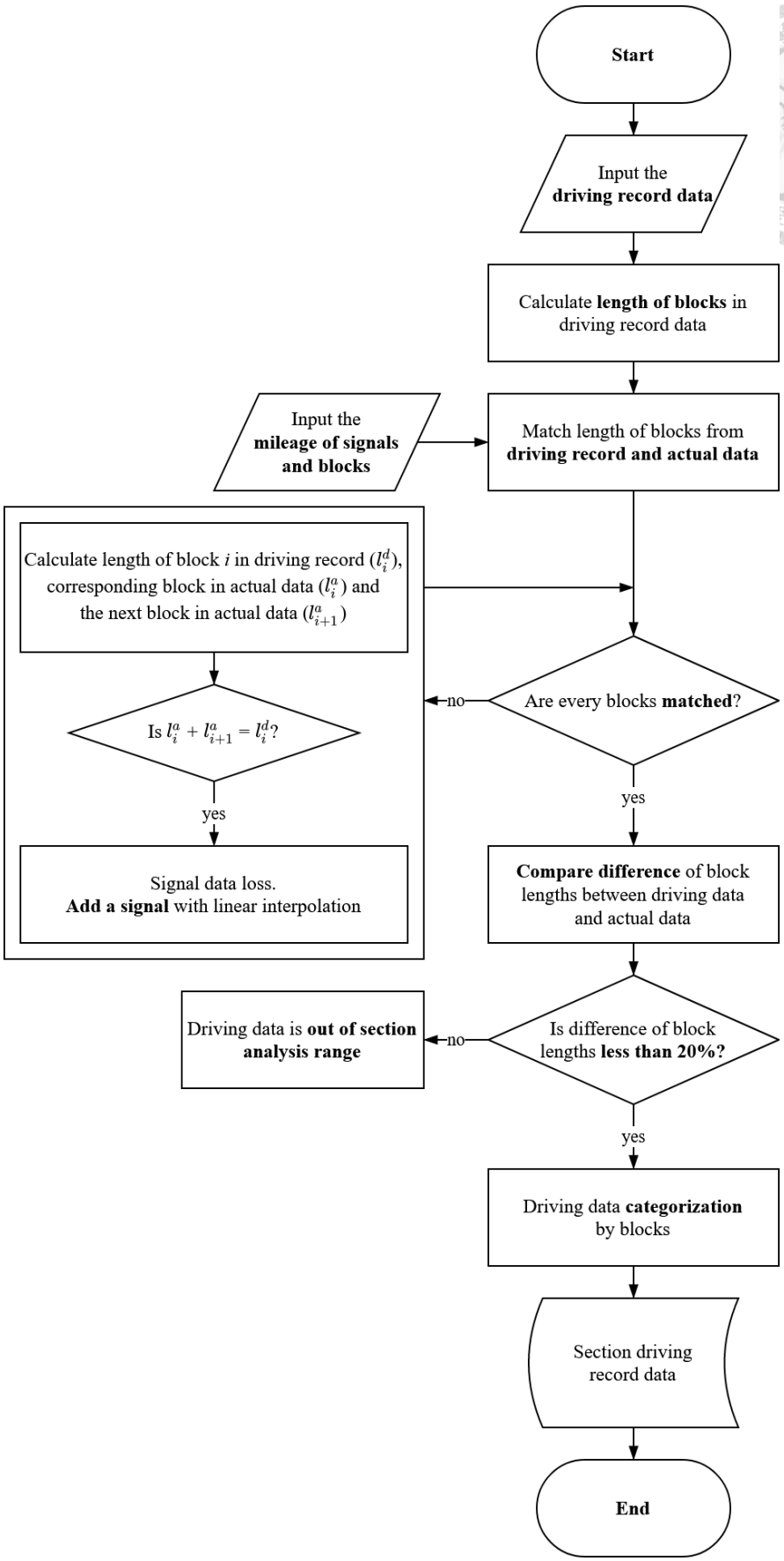


Figure 3-8: Flowchart of section high-risk driving behaviors analysis module



doi:10.6342/NTU202001876

*Step 1: Calculate lengths of every block*

Since locations of signals are accessible in the data, this research decides to categorize data by blocks for section analysis. The concept of categorization by section is to compare the lengths of blocks in the driving record data with those in the actual data. As a result, the first step of this process is to calculate the lengths of blocks (the distance between two signals).

*Step 2: Match block lengths of the driving record data and the actual data*

The second step of this process is to fill in the lost signals because sometimes balises are not detected by ATP. Let the length of block *i* in the driving record data be $l_i^d$, the length of block *i* in actual data be $l_i^a$, and then check if $l_i^d = l_i^a + l_{i+1}^a$.

*Step 3: Interpolate the missing signal*

If $l_i^d = l_i^a + l_{i+1}^a$, a missing signal is within block *i*. Add a signal in block *i* with linear interpolation.

*Step 4: Compare the difference of block lengths between the driving record data and the actual data*

After interpolating the missing signals, compare the new list of block lengths of the driving records to that of actual data again. If less than 80% of the block distances are close to the actual block distances, filter out this driving record because it is out of the analysis scope.

*Step 5: Data categorization*

Finally, divide the driving record data by the signal positions. Categorize the data by the number of blocks for further analysis.



doi:10.6342/NTU202001876

### 3.5 High-risk driving behaviors Analysis Module

In accordance with safety concerns described in the previous sections, this research proposes some modules to identify high-risk driving behaviors based on driving records accessed from ATP. In this research, six types of high-risk driving behaviors are identified and discussed: “operational overspeed,” “ATP service brake,” “ATP emergency brake,” “approach signal overspeed,” “switch signal overspeed,” and “deceleration after target indication.” Attributes and logic of identification will be described in the following sections.

Overspeed is the most intuitive in terms of high-risk driving behaviors of railway drivers. Overspeed can be categorized into overspeed of ATP and overspeed of signals, representing train speed exceeding the speed restriction of ATP and signal, respectively. The overspeed of ATP includes “operational overspeed,” “ATP service brake,” and “ATP emergency brake.” These behaviors are classified in accordance with the severity of the speed limit violation. For operational overspeed, train speed is between the permitted speed profile and ATP service brake profile, while ATP either has no reaction or warns the driver. When train speed exceeds the ATP service brake profile, service brake is activated to slow down the train, and the driver can relieve the brake when the speed is slower than the permitted speed. Once the train speed exceeds the ATP emergency brake profile, the emergency brake is activated immediately and not until the train comes to a full stop can the driver relieve the brake.

In general cases, the speed limit of ATP and signals are identical, except for two conditions: “approach signals” and “switch signals.” Under these conditions, lower speed limits are set after passing through signals, while ATP calculates a smooth speed limit profile from red signals or turnouts (see Figure 3-1). Despite the efficiency of the ATP speed restriction, it is a violation of regulations because drivers should follow



doi:10.6342/NTU202001876

signals instead of ATP in TRA. Moreover, it can be an indicator of reliance on ATP.

Apart from identifying the violation of rules, this research also attempts to identify leading indicators of safety. The proposed leading indicator is "deceleration after target indication," representing driver slowing down the train speed after ATP had sent target indication. Although this action does not violate TRA's regulations, it is not a safe behavior because it is considerably late to decelerate after ATP had sent the target indication. Speed profiles calculated by ATP are close to the limitations of train performance, while drivers should drive trains smoothly in practical operation. Moreover, deceleration after target indication can be treated as an indicator of ATP reliance considering that these drivers probably slow down the train after receiving ATP target indication.

### 3.5.1 Deceleration after Target Indication

Deceleration after target indication indicates that the train driver moves the master controller from accelerating or constant speed to decelerating after hearing the warning from the ATP, as depicted in Figure 3-9. This action can be an indicator of driver reliance on ATP, despite not violating operation regulations in TRA. The speed profile of the target indication is close to the speed restriction profile. Hence, drivers should start deceleration earlier than the target indication. Another concern is on reliance on ATP. Evidently, every important decision and judgement can be done by ATP and easily accessible for drivers. Given that ATP provides a considerably easy way for driving and not violating regulations, some unscrupulous drivers may inevitably manipulate trains relying on ATP. In this research, deceleration after target indication is considered an indicator of reliance on ATP because drivers slow down the train significantly earlier than the target indication in general cases. Consequently, the current research defines deceleration after target indication as a type of high-risk driving behavior.



doi:10.6342/NTU202001876

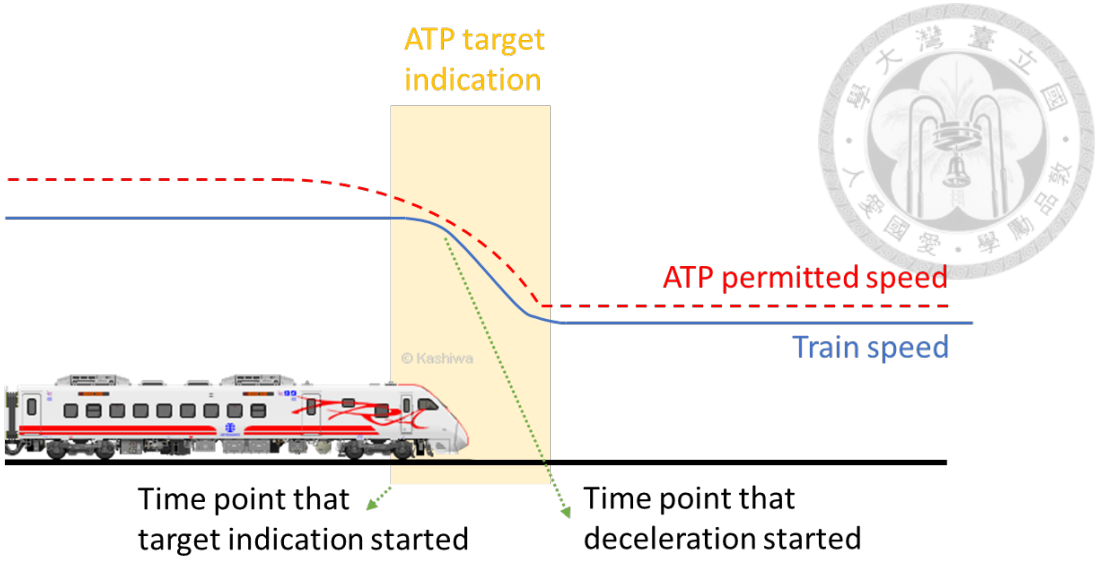


Figure 3-9: Schematic of deceleration after target indication

In deceleration after target indication identification, Theil–Sen regression, which is a method for robust simple linear regression, is used to fit two trend lines. This estimation approach is named after Theil (1950) and Sen (1968). Theil–Sen regression is insensitive to outliers and easy to calculate (Wilcox, 2010). The formula to calculate slope of the Theil–Sen regression model is shown in equation (3):

$$m = median\left(\frac{y_j - y_i}{x_j - x_i}\right) \tag{3}$$

where $(x_i,y_i)$ and $(x_j,y_j)$ are all pairs of sample points in a two-dimensional space. With slope $m$ determined, the y-intercept $b$ is calculated thereafter using the equation (4):

$$b = median\left(y_i - mx_i\right) \tag{4}$$

where $(x_i,y_i)$ represents every sample point in the space. Lastly, the regression function can be expressed as equation (5):

$$y = mx + b \tag{5}$$



doi:10.6342/NTU202001876

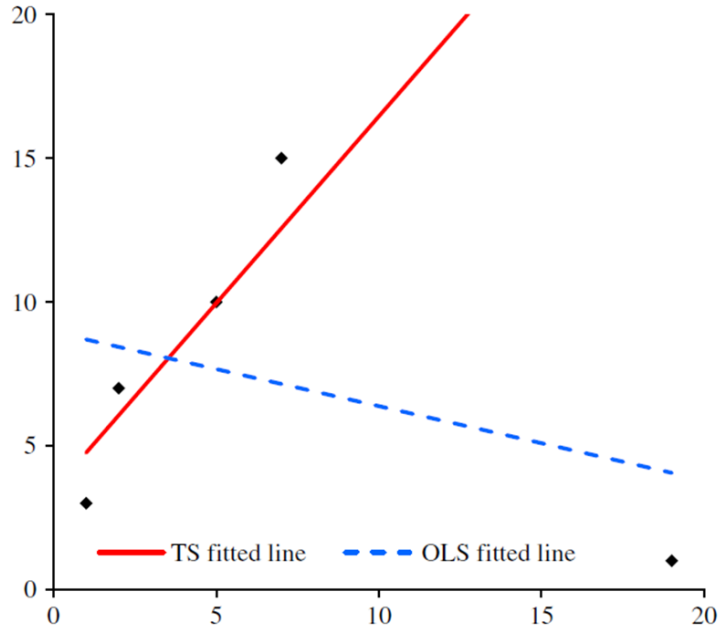


Figure 3-10: Comparison of Theil–Sen regression (TS) and ordinary least squares regression (OLS)

Figure 3-10 shows the comparison between the Theil–Sen regression and ordinary least squares regression (Ohlson and Kim, 2015). Such robust regression technique is applied in this module to develop the trend lines, in order to avoid the effects of outliers. Figure 3-11 demonstrates an example of deceleration time identification. In this figure, the blue points are the train speed before ATP target indication started; the orange points are the train speed after ATP target indication started; and the blue line and orange line respectively represents the trend lines of train speed before and after ATP target indication started. The time of the intersection of the two straight lines is defined as the time point that the train started to decelerate. Once the time of the intersection is later than the time that ATP target indication started, deceleration after target indication is then identified accordingly. Two exceptions are considered in the process. The first one is that the behaviors will not be identified as a high-risk driving behavior if time of the intersection train speed is lower than 25 km/h because it is acceptable to stop and then pass the permissive stop signals under 25 km/h in TRA. The second consideration is to



doi:10.6342/NTU202001876

prevent false judgement in the process. To ascertain that the action is under the beginning of deceleration, two more criteria are added: the acceleration of "before target indication" must not be negative, while the acceleration of "after target indication" must be negative.

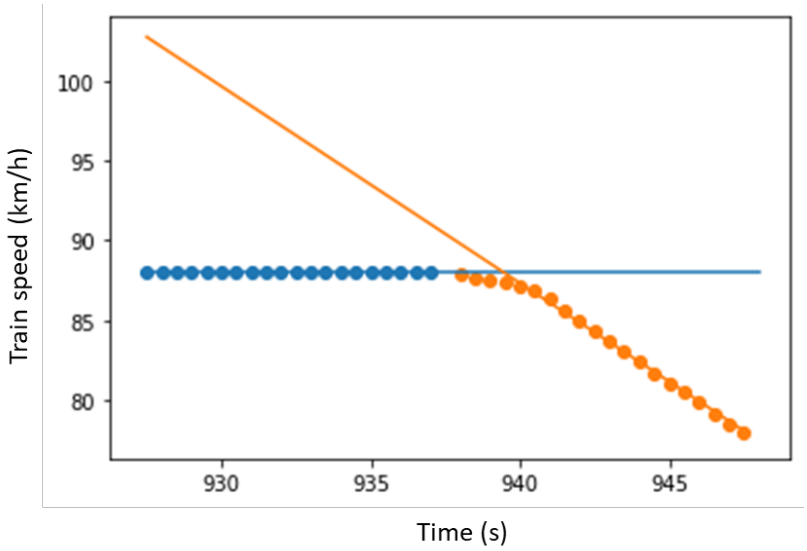


Figure 3-11: An example of deceleration time identification

Deceleration after target indication is checked intermittently. This process is performed for all target indications. The total number of deceleration after target indications is collected as an indicator of high-risk driving behavior for further analysis. The number of deceleration after target indications per kilometer is considered an indicator of deceleration after target indication.

Deceleration after target indication identification is achieved via Scikit-learn, a Python library (Pedregosa et al., 2011).

In this research, the definition of deceleration after target indication is that the point of acceleration change is later than the time target indication starts. Figure 3-12 shows the flowchart of the identification process. The detailed process is shown as follows:



doi:10.6342/NTU202001876

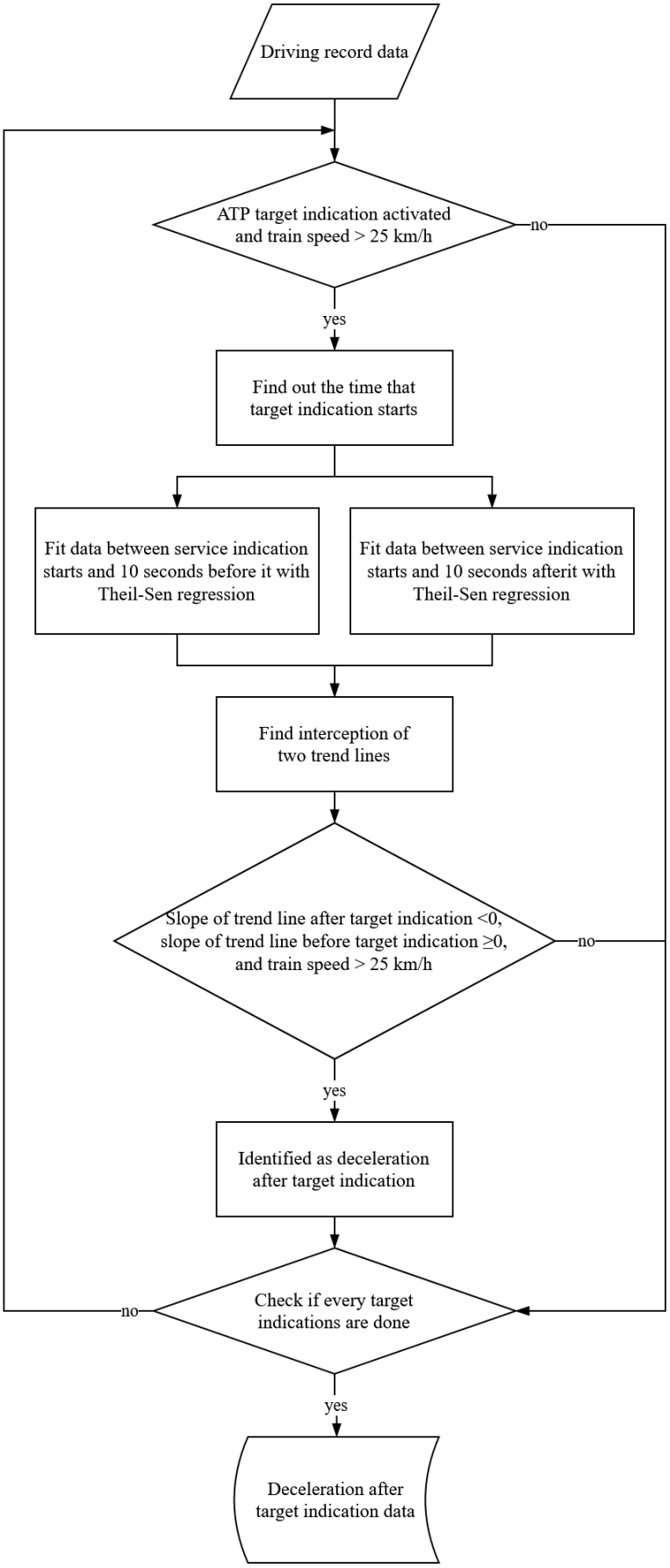


Figure 3-12: Flowchart of deceleration after target indication identification



doi:10.6342/NTU202001876

*Step 1: Identify start time of the target indication*

To identify deceleration after target indication, the first step is to iterate through every target indication and determine the start time of each.

*Step 2: Obtain data before and after target indication starts*

After determining the start time of the target indication, two sections of data are obtained: data between start time of target indication and 10 seconds before it, and data between start time of target indication and 10 seconds after it. The former section is defined as "Before target indication starts," while the latter section is "After target indication starts."

*Step 3: Make two regression lines before and after the start time of the target indication*

Make two Theil–Sen regression lines for before and after target indication starts, and separately determining two trend lines representing before and after target indication starts.

*Step 4: Determine the interception of trend lines for identification*

Find the intersection point of two trend lines, which represent the time point of acceleration change. If the start time of the target indication is earlier than acceleration change, then it is identified as deceleration after target indication.

### 3.5.2 Approach Signal Overspeed

Owing to the different concept of speed limit of ATP and wayside signal, a gap in speed limit exists between them while confronting the approach signals. For wayside signal, a stair-shaped profile of speed limit is used. That is, an approach signal with speed limit of 60 km/h is in advance of a stop signal to ensure that trains can stop before it. However, speed restriction for ATP is the speed that the train will not pass stop signal, as shown in Figure 3-13.



doi:10.6342/NTU202001876

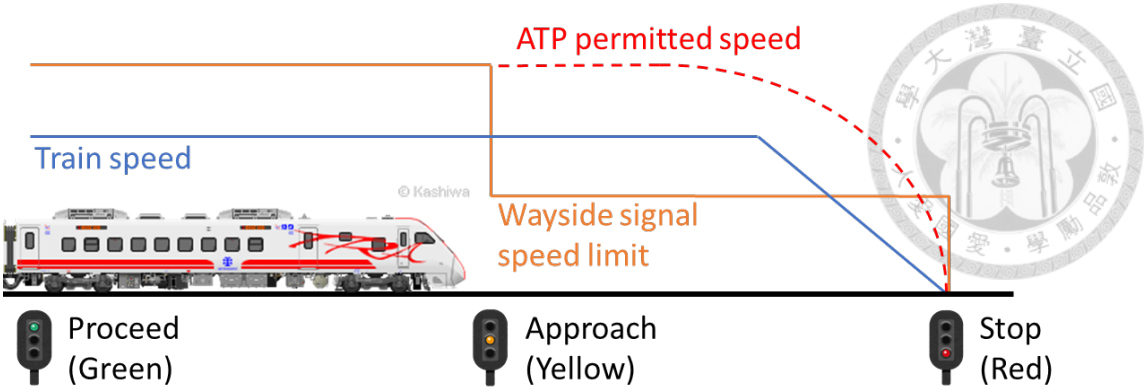


Figure 3-13: Schematic of approach signal overspeed

Owing to the gap of speed restriction between wayside signal and ATP, approach signal overspeed can be an indicator of reliance on ATP. If a driver operating a train relies on ATP, then he may be unaware of the wayside approach signal and overspeed, which is also a violation of regulation in TRA.

The definition of approach signal overspeed is that a train speed exceeds 60 km/h during sections of approach signal. The operation regulations of TRA document an exception: it is permissible to pass the approach signal exceeding 60 km/h during deceleration if there is no sufficient sight distance for the block signal (TRA, 2016). Nevertheless, this condition is not considered in the current research because sight distance data are inaccessible. Moreover, the majority of the signals with insufficient sight distance are complemented with approach signals.

Approach signal overspeed is checked continuously. This process is executed for all approach signals. The total time of approach signal overspeed in seconds per kilometer and the severity rate of approach signal overspeed are considered indicators of the approach signal overspeed. The severity of approach signal overspeed is defined as cumulative sum of the excessive train speed over 60 km/h times the time of approach signal overspeed (see equation (6)), where $t_i$ is the unit time interval, and $v_i$ is the train speed. The unit of severity of approach signal overspeed is $s \cdot km/h$. The severity rate



doi:10.6342/NTU202001876

of approach signal overspeed is the severity of approach signal overspeed divided by driving distance, and the unit of severity rate of approach signal overspeed is $\frac{s \cdot km/h}{train\text{-}km}$.

Figure 3-14 shows the process of approach signal overspeed. The detailed process is shown as follows:

$$\sum_i t_i \cdot (v_i - 60), \quad \forall i \text{ identified as approach signal overspeed} \tag{6}$$

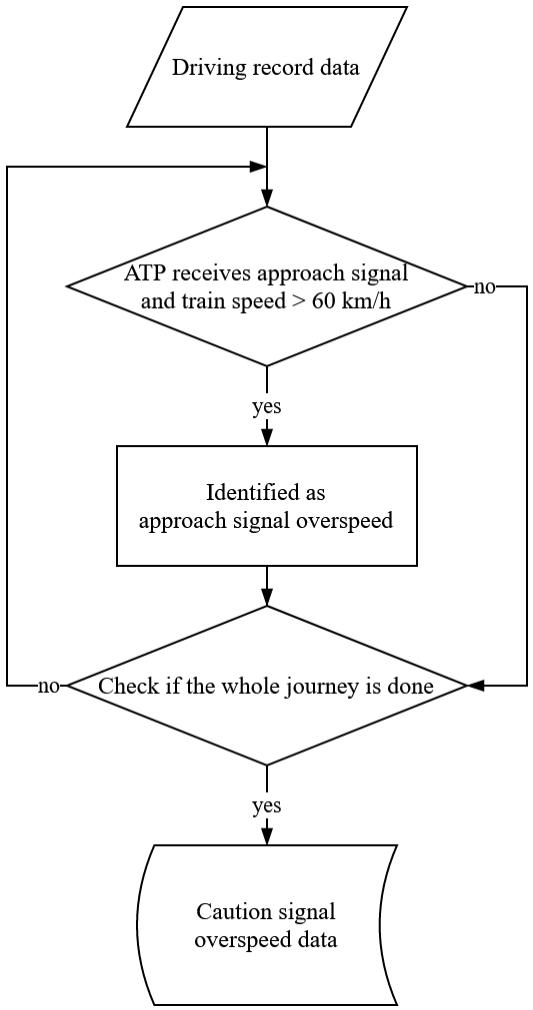


Figure 3-14: Flowchart of approach signal overspeed identification

*Step 1: Identify the approach signals*

In order to identify the approach signal overspeed, the first step is to iterate through every approach signal, and to find out the train speed while confronting approach signals.



doi:10.6342/NTU202001876

*Step 2: Check if overspeeding*

If the approach signal and the train speed higher than 60 km/h are detected simultaneously, it is identified as approach signal overspeed at that time. Cumulative time of approach signal overspeed and cumulative sum of the excessive train speed over the ATP speed restriction times the time of approach signal overspeed are calculated as indicators of high-risk driving behavior.

### 3.5.3 Switch Signal Overspeed

Four types of switch signals in TRA have speed limits of 25 km/h, 35 km/h, 45 km/h, and 60 km/h. When passing a switch signal, the train has to slow down to pass through a turnout. Similar to the approach signal overspeed, a gap between ATP and wayside signal also exists for switch signals. The reason to set the switch signal apart from the approach signal is that the stop points of former are turnouts.

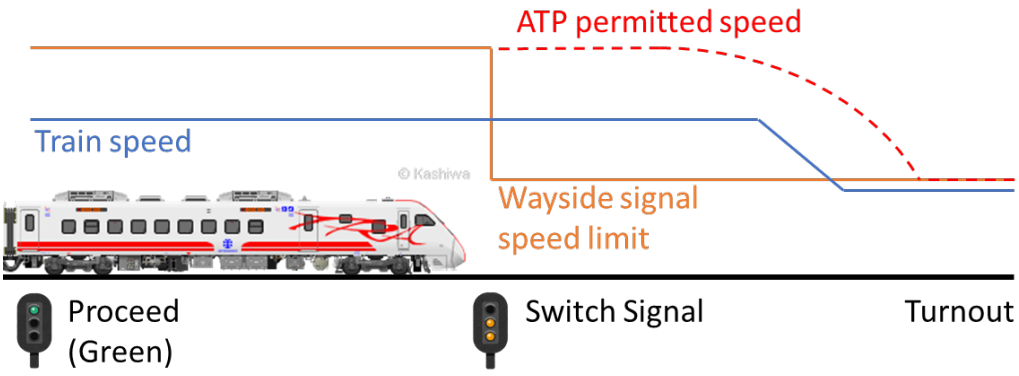


Figure 3-15: Schematic of former switch signal overspeed

The definition of switch signal overspeed indicates that train speed exceeds the speed limit of each switch signals. Unlike approach signal overspeed, switch signal overspeed is checked intermittently, which is a result of the special speed limit of turnouts. Switch signals restrict train speed before and during trains passing through a turnout, while speed limit resumes to 60 km/h after passing through a turnout. Given that the positions of turnouts are inaccessible, switch signal overspeed is identified



doi:10.6342/NTU202001876

intermittently. This process is performed for all target indications. The total number of switch signal overspeeds per kilometer and the severity rate of switch signal overspeed are considered indicators of switch signal overspeed. The severity of switch signal overspeed is defined as cumulative sum of the excessive train speed over the speed limit of the switch signal (see equation (7)), where $v_i$ is the train speed, and $s_i$ is the speed limit of the switch signal. The unit of severity of switch signal overspeed is km/h. The severity rate of switch signal overspeed is the severity of switch signal overspeed divided by driving distance, and the unit of severity rate of approach signal overspeed is $\frac{km/h}{train\text{-}km}$. Figure 3-16 shows the process of switch signal overspeed. The detailed process is shown as follows:

$$\sum_i (v_i - s_i), \quad \forall i \text{ identified as switch signal overspeed} \tag{7}$$



doi:10.6342/NTU202001876

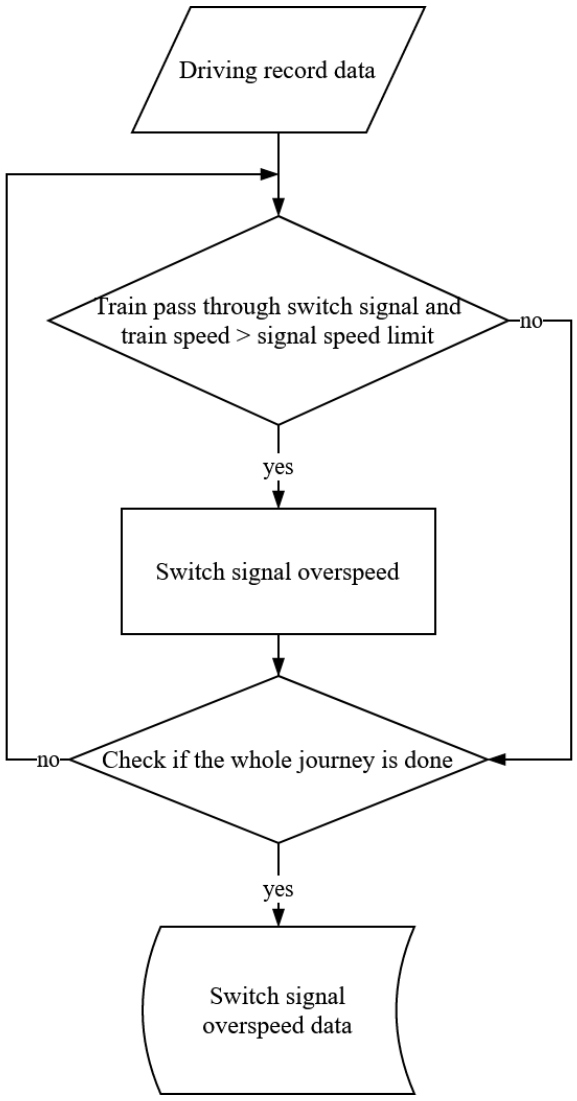


Figure 3-16: Flowchart of switch signal overspeed identification

*Step 1: Identify the switch signals*

In order to identify the switch signal overspeed, the first step is to iterate through every switch signal location, and to find out the train speed while confronting switch signals.

*Step 2: Check if overspeeding*

If the train passes through a switch signal and the train speed exceeds the corresponding speed limit of the switch signal simultaneously, it is identified as switch signal overspeed. The number of switch signal overspeed and cumulative sum of the excessive train speed over the speed limit of the switch signal are calculated as indicators of high-risk driving behavior.



doi:10.6342/NTU202001876

### 3.5.4 Operational Overspeed

Operational overspeed indicates that train speed exceeds the speed limit calculated by ATP. In TRA's regulation, a 3-km/h leeway is provided for overspeeding and has been included in ATP. However, train speed less than 25 km/h is excluded.

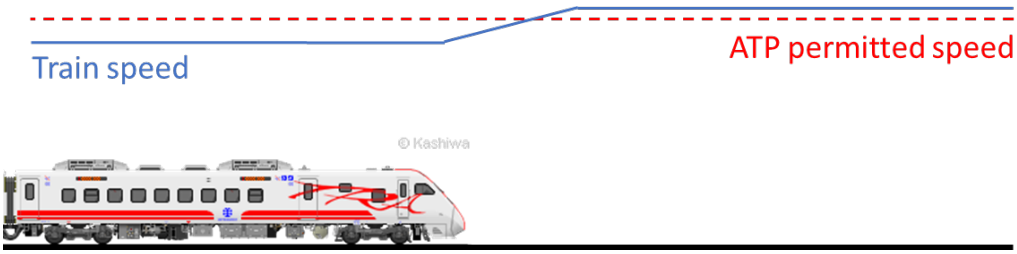


Figure 3-17: Schematic of operational overspeed

The definition of operational overspeed indicates that train speed exceeds ATP permitted speed during train operation. Operational overspeed is checked continuously, and this process is executed throughout the journey. The total time of operational overspeed in seconds per kilometer and the severity rate of operational overspeed are considered indicators of operational overspeed. The severity of operational overspeed is defined as cumulative sum of the excessive train speed over the ATP permitted speed times the time of operational overspeed (see equation (8)), where $t_i$ is the unit time interval, $v_i$ is the train speed, and $r_i$ is the ATP permitted speed. The unit of severity of operational overspeed is $s \cdot km/h$. The severity rate of operational overspeed is the severity of operational overspeed divided by driving distance, and the unit of severity rate of operational overspeed is $\frac{s \cdot km/h}{train\text{-}km}$. Figure 3-18 shows the process of operational overspeed. The detailed process is shown as follows:

$$\sum_i t_i \cdot (v_i - r_i), \quad \forall i \text{ identified as operational overspeed} \tag{8}$$



doi:10.6342/NTU202001876

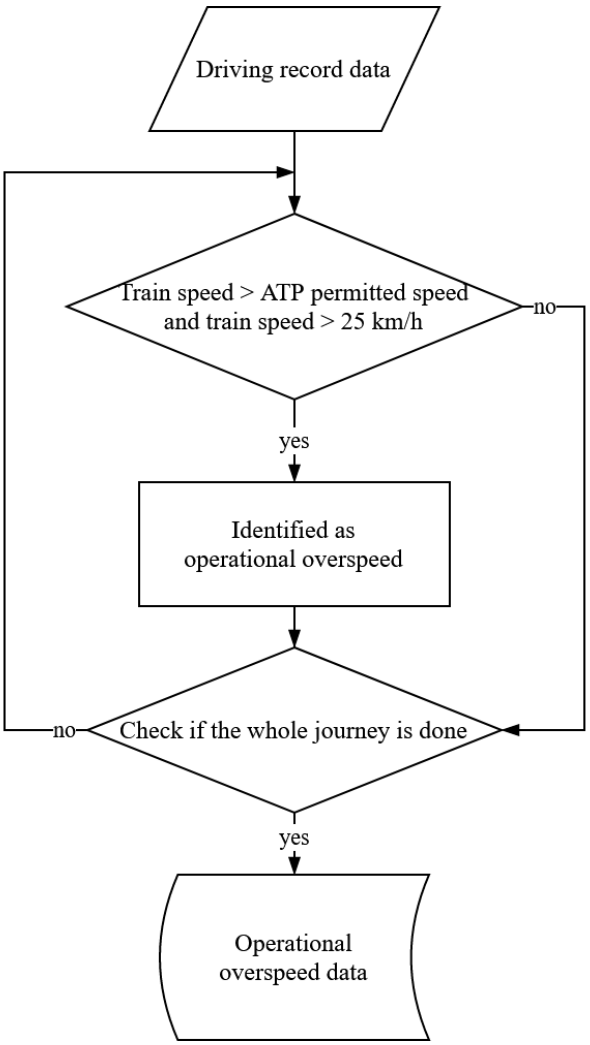


Figure 3-18: Flowchart of operational overspeed identification

*Step: Identify operational overspeed*

The process for operational overspeed identification is to iterate along the whole journey. If train speed exceeds ATP permitted speed and over 25 km/h, operational overspeed is identified. Cumulative time of operational overspeed and cumulative sum of the excessive train speed over the ATP permitted speed times the time of operational overspeed are calculated as indicators of high-risk driving behavior.

### 3.5.5 ATP Service Brake

ATP service brake means that the maximum service brake is activated by ATP. Once train speed exceeds the service brake profile calculated by ATP, the maximum



doi:10.6342/NTU202001876

service brake would be activated automatically to lower the train speed. Not until train speed is slower than ATP speed limit could the driver release the ATP service brake.

The definition of ATP service brake is that there exists a maximum service brake record triggered by ATP. ATP service brake is checked continuously, which is executed throughout the entire journey. The total number of ATP service brakes per kilometer is considered an indicator of ATP service brake. Figure 3-19 shows the process of ATP service brake. The detailed process is shown as follows:

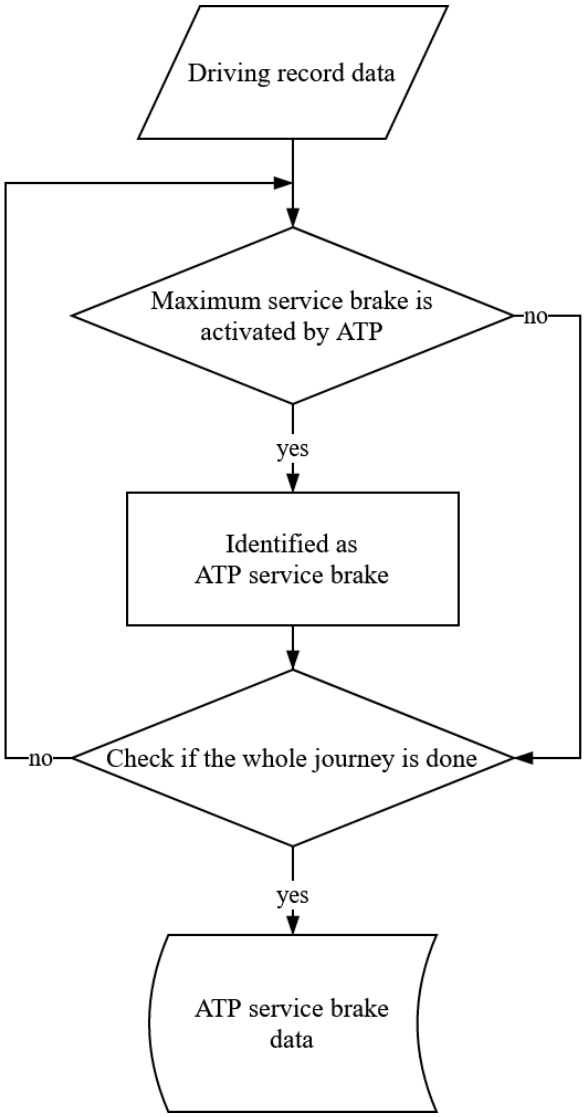


Figure 3-19: Flowchart of ATP service brake identification



doi:10.6342/NTU202001876

*Step: Identify ATP service brake*

The process for ATP service brake identification is to iterate along the whole journey. ATP service brake is identified by brake activation record. The number of ATP service brakes is calculated as an indicator of high-risk driving behavior.

### 3.5.6 ATP Emergency Brake

ATP emergency brake means that emergency brake is activated by ATP. Once train speed exceeds the emergency brake profile that is calculated by ATP, emergency brake would be activated automatically to stop the train. Not until the train had stopped could the driver release the ATP emergency brake.

The definition of ATP emergency brake is that there exists an emergency brake record triggered by ATP. ATP emergency brake is checked continuously, and this process is executed throughout the entire journey. The total number of ATP emergency brakes per kilometer is considered an indicator of ATP emergency brake. Figure 3-20 shows the process of ATP emergency brake. The detailed process is shown as follows:



doi:10.6342/NTU202001876

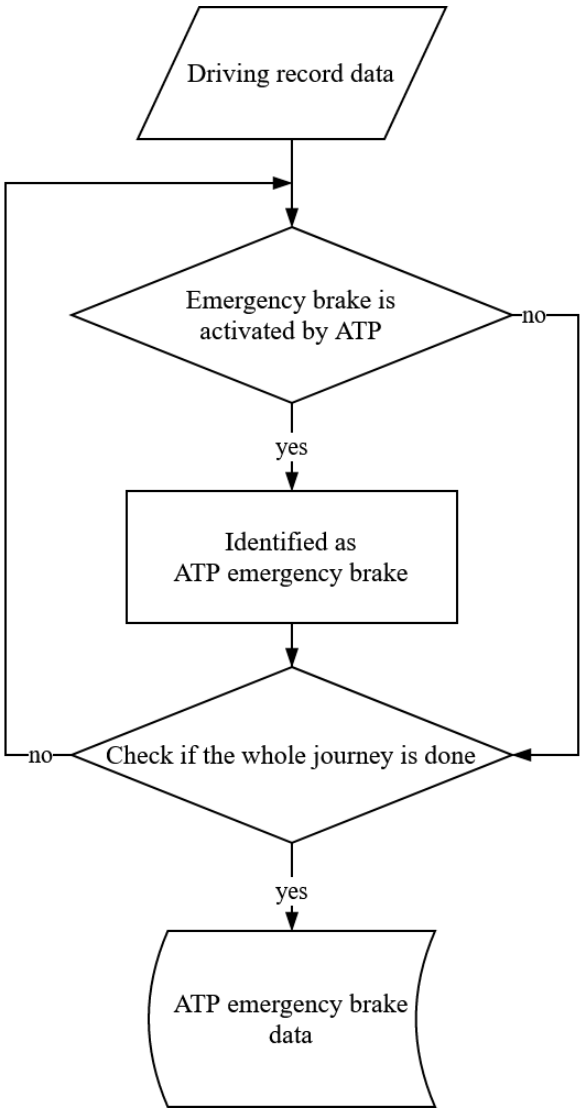


Figure 3-20: Flowchart of ATP emergency brake identification

*Step: Identify ATP service brake*

The process for ATP emergency brake identification is to iterate along the whole journey. ATP emergency brake is identified by brake activation record. The number of ATP emergency brakes is calculated as an indicator of high-risk driving behavior.

## 3.6 Integrated Risk Index for Driving Behaviors (IRIDB) Evaluation Module

In section 3.5, six indicators of high-risk driving behaviors are obtained. However, it is difficult to assess high-risk drivers or sections with these six indicators because they are not comparable. This research proposes an integrated risk index for driving



doi:10.6342/NTU202001876

behaviors (IRIDB) combining the six indicators and then generates a score of risk that can be applied to compare risk between different drivers or sections. Weighted mean is used to determine the risk index. IRIDB evaluation module will be discussed in the forthcoming paragraphs.

### 3.6.1 Normalization of High-risk Driving Behavior Indicators

In section 3.5, six high-risk driving behavior indicators are calculated for risk assessment. However, the units and scales of indicators vary with each high-risk driving behavior. To integrate all indicators into a single index normalization is a necessary process. Given that every indicator ranges from 0 to ∞, this research normalizes these indicators by dividing them by the expected value of each indicator separately. The normalization is calculated using equation (9) and (10).

$$r_i^d = \frac{\text{indicator of high-risk behavior } i \text{ for driver } d}{\text{expectancy of indicator of high-risk behavior } i}, \quad \forall i \in I,\ d \in D, \qquad (9)$$

$$r_i^s = \frac{\text{indicator of high-risk behavior } i \text{ for section } s}{\text{expectancy of indicator of high-risk behavior } i}, \quad \forall i \in I,\ s \in S. \qquad (10)$$

The evaluation of risk index is categorized into two parts: analyses of high-risk driver and high-risk section. In equation (9) and (10), $I$ is a set of six high-risk driving behaviors; $D$ is a set of drivers; $S$ is set of sections; $r_i^d$ and $r_i^s$ are the normalized indicators of high-risk driving behavior $i$ for driver $d$ and section $s$, respectively; $r_i^d$ and $r_i^s = 1$ represents that performance is identical to the overall performance. If $r_i^d$ or $r_i^s > 1$, then the performance is below average; while $r_i^d$ or $r_i^s < 1$ representing performance is above average.



doi:10.6342/NTU202001876

### 3.6.2 Determination of Weightings

To determine the weighting of each high-risk driving behaviors, this research applies analytic hierarchy process (AHP) as the methodology for evaluation. AHP is a structured technique for decision-making in a complex system. Since it was first developed in the 1970s, AHP has been applied to solve different problems, such as selection, evaluation, planning, ranking, and decision making. AHP has been used in a variety of fields, such as economics, management, finance, industry, and transportation (Vaidya and Kumar, 2006; Vargas, 1990; Zahedi, 1986). Moreover, AHP is applicable to measure risks between hazard events. Moreover, AHP is applicable to measure risks between hazard events. Raviv et al. (2017) used AHP to determine the weighting of different near-miss events and accidents in the proposed risk assessment model.

The purpose of AHP is to transfer a complex system into a hierarchy structure, decomposing it in different layers. Thereafter, pairwise comparisons are used to measure priorities or importance of each factors in a layer. Lastly, a quantified evaluation of the overall system can be obtained (R. W. Saaty, 1987; T. L. Saaty, 1988, 1990).

The evaluation is based on the concept of pairwise comparisons. AHP uses ratio between factors to evaluate their weighting. If there are *n* factors to be analyzed in a layer, then $C_2^n$ pairwise comparisons are needed. The scale of pairwise comparison is listed in Table 3-3 (T. L. Saaty, 2008).



doi:10.6342/NTU202001876

Table 3-3: Scale of pairwise comparisons in AHP

| Intensity of importance | Definition |
|---|---|
| 1 | Equal importance. |
| 3 | Moderate importance. |
| 5 | Strong importance. |
| 7 | Very strong importance. |
| 9 | Extreme importance. |
| 2, 4, 6, 8 | Intermediate between two neighboring scales. |
| Reciprocals of above | If *i* has one of intensity of importance compared to *j*, then *j* has the reciprocal value of intensity of importance compared to *i*. |

The result of pairwise comparisons can form an $n \times n$ positive reciprocal matrix $A$. the elements $a_{ij}$ in $A$ are the intensity of the importance values of $i$ compared with $j$ for all $i, j = 1, \ldots, n$. The reciprocal characteristic provides that $a_{ij} = 1/a_{ij}$ for all $i, j = 1, \ldots, n$, as shown in equation (11).

$$A = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ a_{21} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{bmatrix} = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ \frac{1}{a_{12}} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{a_{1n}} & \frac{1}{a_{2n}} & \cdots & 1 \end{bmatrix} \tag{11}$$

The relative importance (weighting) between different factors can be obtained by calculating the maximum eigenvalue $\lambda_{max}$ and the corresponding eigenvector of matrix $A$. The eigenvector represents the weightings of each factor.

Based on the assumption of AHP, the relative importance in the matrix should fulfill transitivity. A consistency check is often used to ensure transitivity of the evaluation. T. L. Saaty (2000) explained that consistency can be evaluated using



doi:10.6342/NTU202001876

consistency index (C.I.) and consistency ratio (C.R.). Equation (12) and (13) show the calculation of C.I. and C.R.

$$C.I. = \frac{\lambda_{\max} - n}{n-1}, \tag{12}$$

$$C.R. = \frac{C.I.}{R.I.}. \tag{13}$$

The random index (R.I.) in equation (13) is the average C.I. of the randomly generated reciprocal matrices (Forman, 1990). Table 3-4 presents R.I. values for different number of judgements *n*. T. L. Saaty (2000) suggested that it is tolerable if C.R. < 0.1. Given this process, the values of weightings can be obtained to evaluate the severity of high-risk driving behaviors.

Table 3-4: Table of random index (R.I.)

| *n* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| R.I. | 0.00 | 0.00 | 0.52 | 0.89 | 1.11 | 1.25 | 1.35 | 1.40 | 1.45 | 1.49 | 1.51 | 1.54 | 1.56 | 1.57 | 1.58 |

The established structure of AHP is presented in Figure 3-21. Based on the structure, a questionnaire with the concept of pairwise comparisons is made for assessment. The research process of AHP is presented in Figure 3-22.



doi:10.6342/NTU202001876

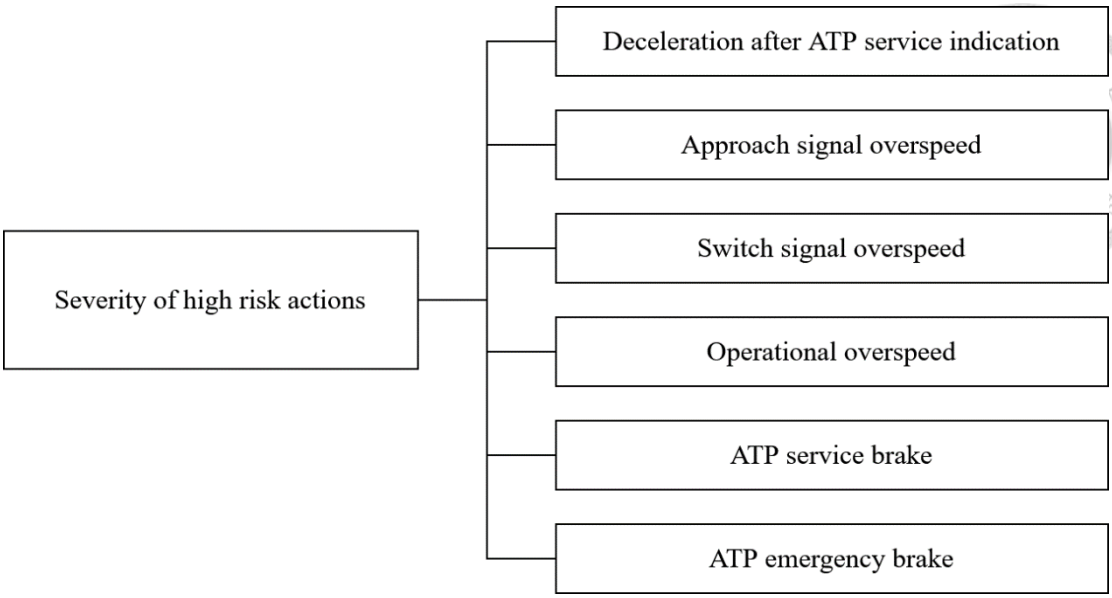


Figure 3-21: Hierarchy structure of the high-risk driving behavior severity analysis

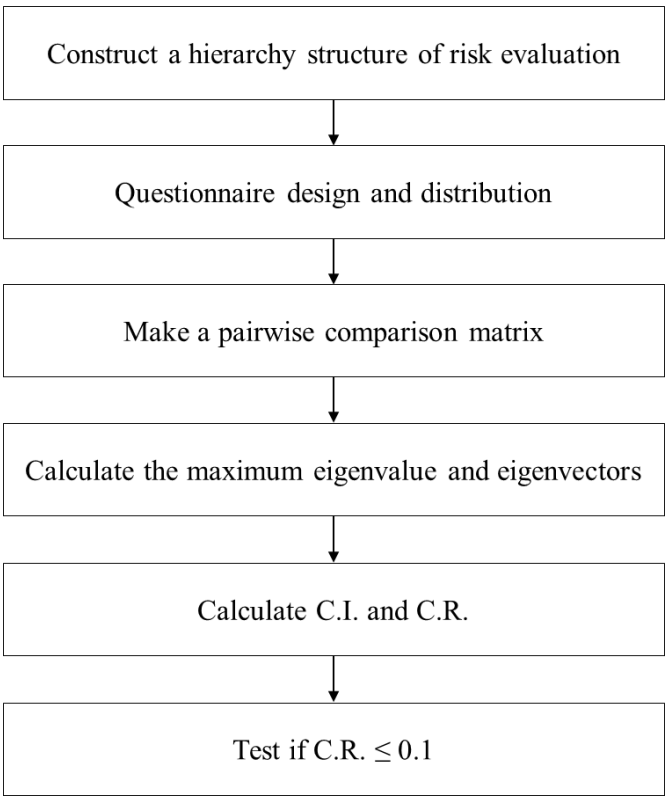


Figure 3-22: Flowchart of AHP in this research

Based on the established hierarchy structure, this research conducts an expert survey to evaluate severity of the six high-risk driving behaviors. The participants in this survey are experts from academic, practical, and supervision units. This survey was conducted via questionnaire. The extensive knowledge and experience of these experts



doi:10.6342/NTU202001876

form the basis for the establishment of an integrated risk index. Once questionnaires were recovered, consistency tests were implemented. Only questionnaires that pass the consistency test would be analyzed.

### 3.6.3 Evaluation of IRIDB

To further assess the overall performance of each driver or in each section, this research also proposes an index to integrate the six high-risk driving behaviors. The method of weighted mean is applied for integrated index calculation, as shown in equation (14) and (15) for driver and section analyses, respectively:

$$IRIDB^d = \sum_i w_i \cdot r_i^d \,, \quad \forall i \in I,\ d \in D, \tag{14}$$

$$IRIDB^s = \sum_i w_i \cdot r_i^s \,, \quad \forall s \in S,\ d \in D. \tag{15}$$

IRIDB of each driver is calculated using the weighted average of the normalized indicators of high-risk driving behaviors. In equation (14), $IRIDB^d$ represents IRIDB for driver $d$; $w_i$ is the weighting of high-risk driving behavior $i$, which is obtained from the AHP questionnaire analyzed in section 3.6.2; and $r_i^d$ is the normalized indicator of high-risk driving behavior $i$ for driver $d$, the calculation method of which is mentioned in section 3.6.1. The IRIDB formula for sections is the same as those for drivers (see equation (15)). The only difference is using the normalized indicator of high-risk driving behavior $i$ for section $s$.

For approach signal overspeed, switch signal overspeed, and operational overspeed, two indicators, the cumulative time and the severity of each high-risk driving behavior, are calculated. Since the distribution of the two indicators for all high-risk driving behaviors are similar, only the cumulative time is used for IRIDB evaluation.



doi:10.6342/NTU202001876

# CHAPTER 4 CASE STUDY

In this chapter, two case studies, one for high-risk driving behaviors and the other for high-risk sections, are performed to demonstrate the application of the proposed analysis framework. At first, Section 4.1 introduces the Puyuma express trains, which is chosen for high-risk driving behaviors analysis and high-risk sections analysis. The sections between Shulin and Hualien, which is chosen for high-risk driving behaviors analysis and high-risk sections analysis, is also introduced. The case study for high-risk driving behaviors analysis is exhibited in Section 4.2. And then the case study for high-risk sections analysis is exhibited in Section 4.3. Finally, Section 4.4 briefly summarizes and discusses the results of these two case studies.

## 4.1 Trainsets and Network of TRA

TRA started operating conventional railway in Taiwan since 1887. The railway network in Taiwan forms a big circular mainline surrounding Taiwan with nine branch lines goes deep into plains and hills.

Until 2018, 256 locomotives and 1,235 multiple-unit trains had been under operation in TRA. The Puyuma express is serviced by TEMU 2000 electrical multiple unit trainset, which is made by Nippon Sharyo. The Puyuma express trains was imported from Japan in 2012 and started operation in 2013, aiming to improve the connection between eastern and western Taiwan. As a tilting train, the Puyuma express trains operates faster than non-tilting trains in curving areas, so the Puyuma express trains are suitable to operate in curving Yilan Line and North-link Line. In July 2019, 12 routine trains were operating in Eastern Main Line (between Shulin and Taitung, one to two extra trains on weekends), two operating in Western Main Line (between Nangang and Chaozhou, has been extended to Fangliao in December 2019), and three cross-line trains every day.



doi:10.6342/NTU202001876

In 2018, there were totally 241 stations and 1,065 km operation mileages in TRA network (TRA, 2019). The network of TRA can roughly be divided into four parts: Western Main Line, Eastern Main Line, South-link Line, and the branch lines. Figure 4-1 shows the section for the case study of high-risk section analysis. This section starts from Shulin and ends at Hualien, composed of three lines. The section between Shulin and Badu is a part of Western Main Line, which is the busiest section in TRA because the trains in both Western Main Line and Eastern Main Line are operating here. The section between Badu and Su'aosin belongs to Yilan Line. The northern half of this section is in a mountainous region, having numerous sharp curves; while the southern half of this section is in a plain region. Xinma station, the location where the Puyuma derailment took place, is also in this section. The section between Su'aosin and Hualien is North-link Line. Although this section crosses a mountain range, long tunnels and bridges make it a relative straight section compared with the other two parts.



doi:10.6342/NTU202001876


Figure 4-1: Analysis network of Puyuma express train

## 4.2 Analysis of High-Risk Driver

This research collects the ATP driving record figures of Puyuma express trains from May 2019 to July 2019. There are 4,586 effective data with 540 drivers and 811,004.7 train-km operation mileage. The driving distance of a single driver ranges between 24.1 train-km and 6708.7 train-km. The average driving distance is 1501.9 train-km and the standard deviation is 1207.1 train-km. Table 4-1 and Figure 4-2 shows basic statistics of the operation mileage of the drivers collected.



doi:10.6342/NTU202001876

Table 4-1: Statistics of driving distance of drivers

| | | |
|---|---|---|
| Average | 1501.9 | train-km |
| Standard deviation | 1207.1 | train-km |
| Minimum | 24.1 | train-km |
| 1st quartile | 622.5 | train-km |
| Median | 1175.8 | train-km |
| 3rd quartile | 1941.0 | train-km |
| Maximum | 6708.7 | train-km |

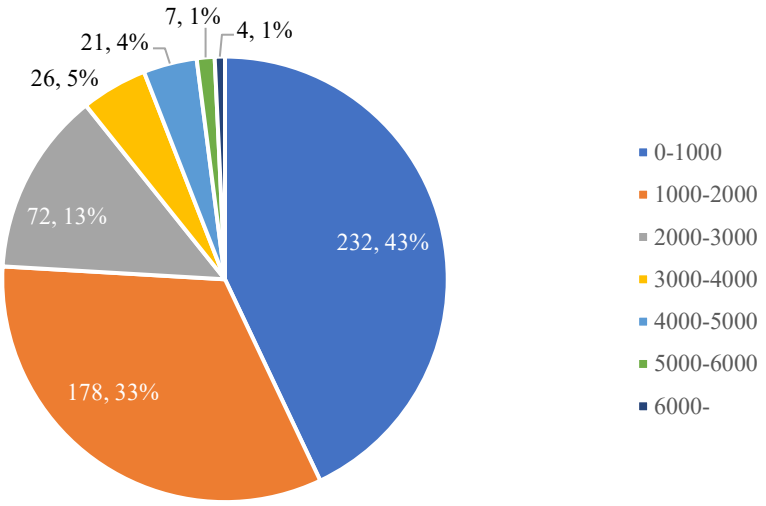


Figure 4-2: Circle graph of the operation mileage of the drivers collected (train-km)

### 4.2.1 Deceleration After Target Indication

Among the 811,004.7 train-km operation distance of the 540 drivers, 5,507 times of decelerations after target indication are committed in total. The overall frequency of deceleration after target indication is $6.79\times10^{-3}$ times per train-km.



doi:10.6342/NTU202001876

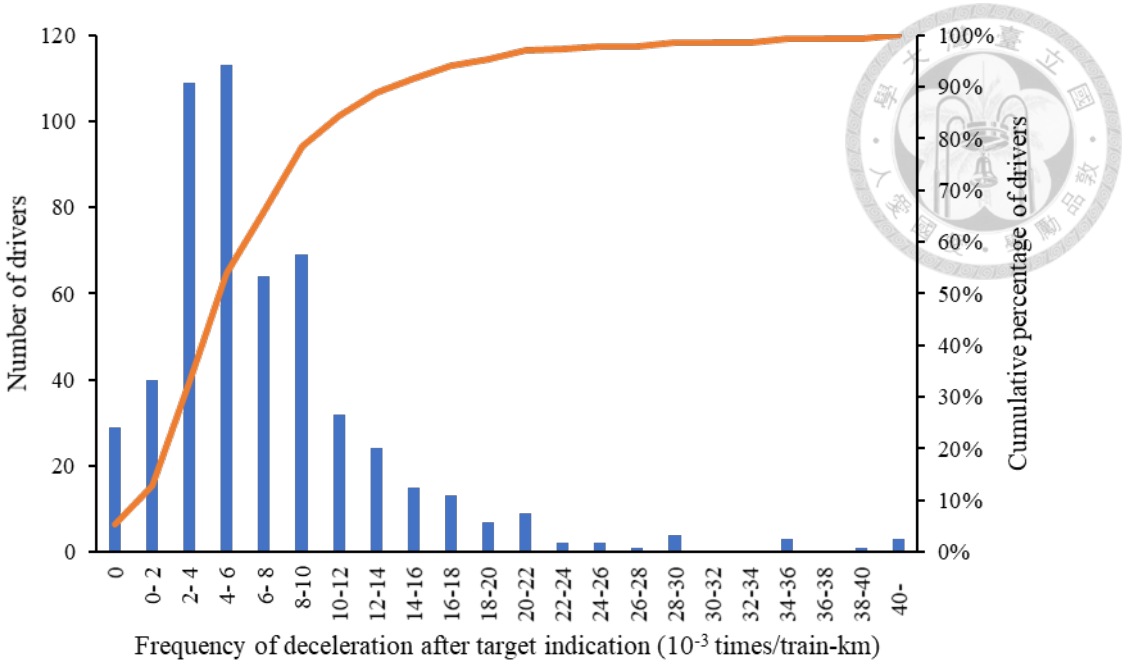


Figure 4-3: Histogram of the frequency of deceleration after target indication

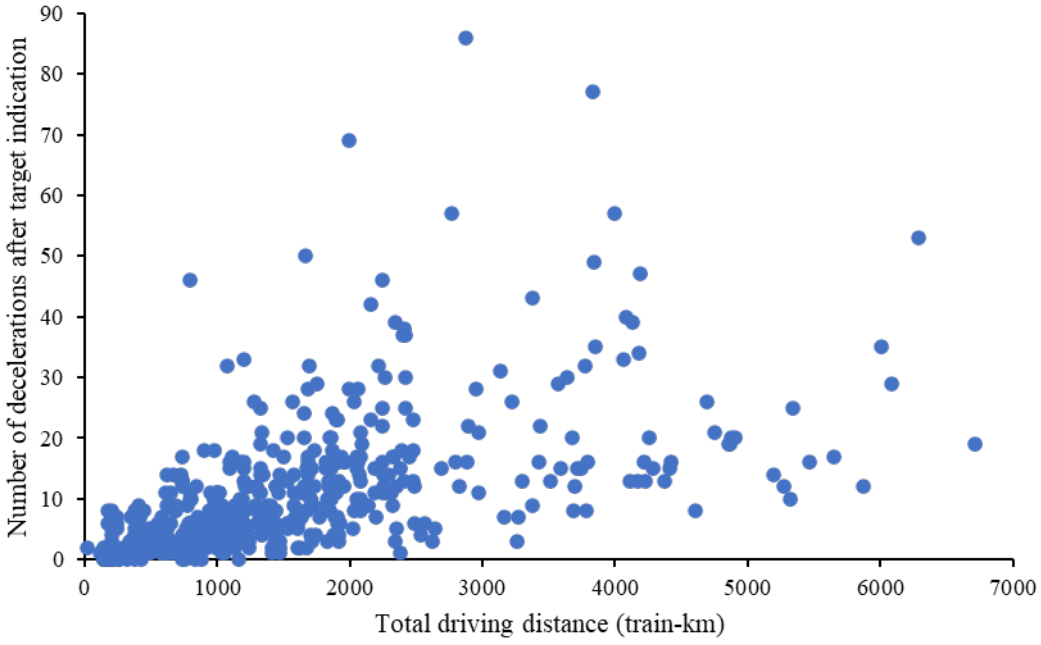


Figure 4-4: Scatter plot of deceleration after ATP warning and driving distance

Figure 4-3 depicts the distribution of frequency of deceleration after target indication. Incidence of deceleration after target indication is discretized by $1\times10^{-3}$ times/train-km for all drivers. This figure shows that 78% of drivers commits deceleration after target indication less than 0.01 times/train-km. On the other hand,



doi:10.6342/NTU202001876

there are some outliers whose frequency of deceleration after target indication are much higher than other drivers. Figure 4-4 shows the relationship between the number of deceleration after target indication and the driving distance for each driver. Although the trend for most drivers is concentrated, there are some extreme drivers with much higher incidence of deceleration after target indication than the average. Moreover, most of these extreme drivers had operated trains for more than 1,000 train-km, which means that the outliers in Figure 4-3 are not the drivers with small samples. With these two figures, it is found that there are some drivers who are accustomed to decelerate after target indication are existing in TRA.

### 4.2.2 Approach Signal Overspeed

Among the 811,004.7 train-km operation distance of the 540 drivers, 72,797 seconds of approach signal overspeed are committed in total. The train speed exceeds the approach signal speed limit by 1.44 km/h averagely. The overall frequency of approach signal overspeed is $8.98\times10^{-2}$ seconds per train-km.

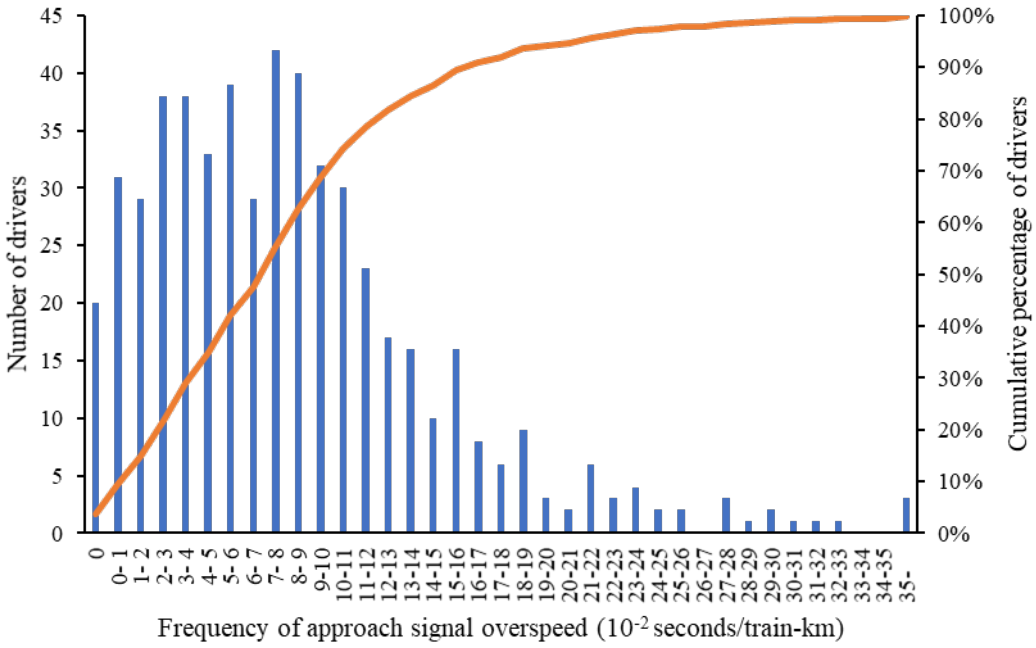


Figure 4-5: Histogram of the frequency of approach signal overspeed



doi:10.6342/NTU202001876

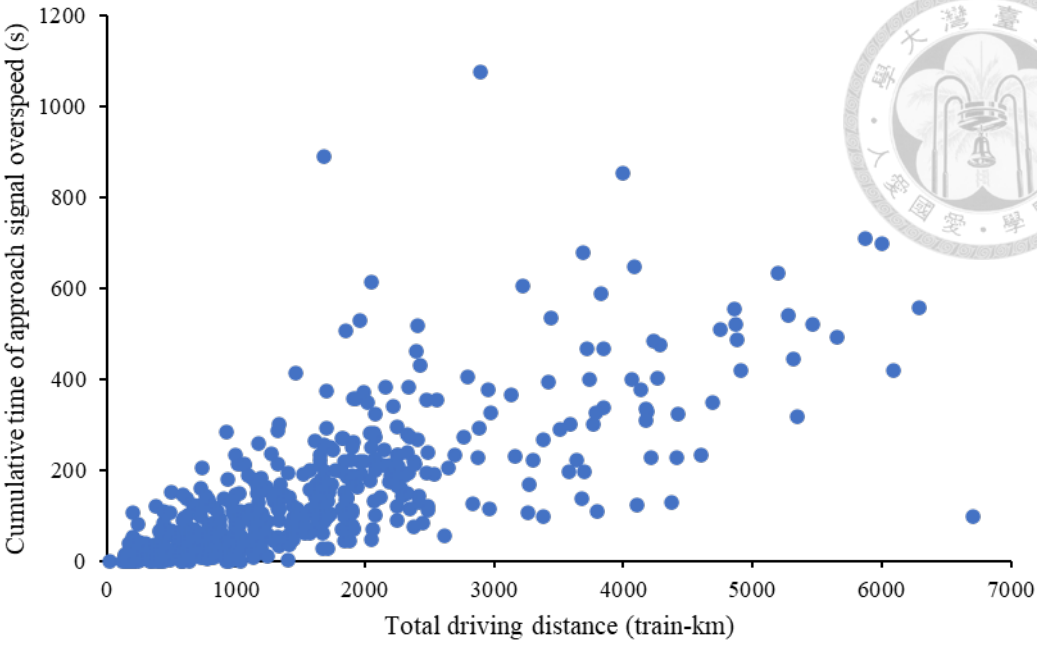


Figure 4-6: Scatter plot of approach signal overspeed and driving distance

Figure 4-5 depicts the distribution of frequency of approach signal overspeed. The incidence of approach signal overspeed is discretized by 0.01 seconds/train-km for all drivers. Similar to deceleration after target indication, there are about 85% drivers whose frequency of approach signal overspeed is less than 0.15 seconds/train-km. Figure 4-6 also illustrates the relationship between the time of approach signal overspeed and the driving distance for each driver. It is noted that the distribution of approach signal overspeed is the most concentrated among the six high-risk driving behaviors. It may represent that the overall performance of drivers on approach signal overspeed is more coherent than the other high-risk driving behaviors.



doi:10.6342/NTU202001876

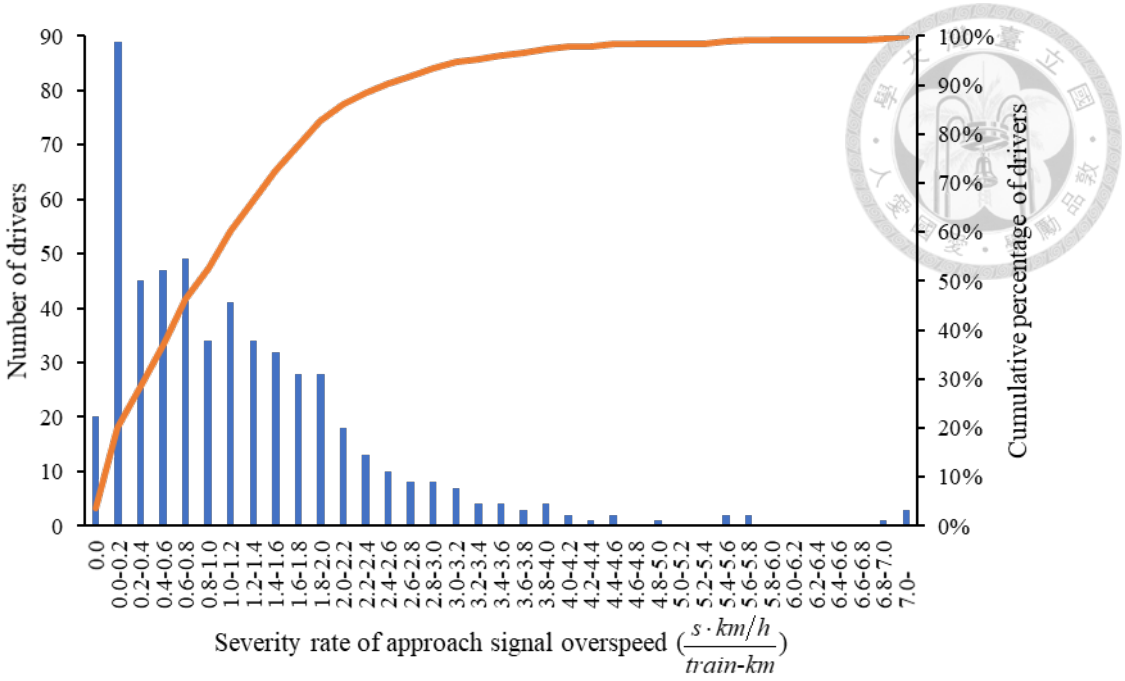


Figure 4-7: Histogram of the severity rate of approach signal overspeed

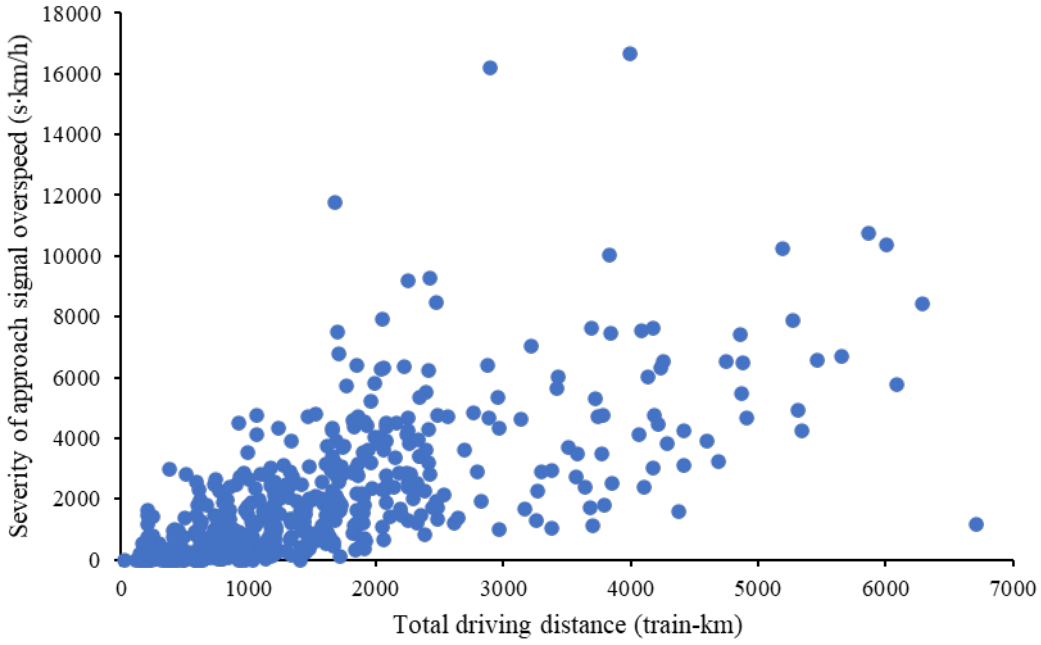


Figure 4-8: Scatter plot of severity of approach signal overspeed and driving distance

Figure 4-7 and Figure 4-8 show the result of severity of approach signal overspeed. According to Figure 4-6 and Figure 4-8, there is no significant difference of distribution between the two figures. It can be inferred that excessive train speed over the speed limit of approach signal does not varies significantly by drivers. For the normal drivers,



doi:10.6342/NTU202001876

excessive train speed over the speed limit of approach signal is quite similar. Only a few drivers who drives relying on ATP have much more severity of approach signal overspeed compared with the others.

### 4.2.3 Switch Signal Overspeed

Among the 811,004.7 train-km operation distance of the 540 drivers, 686 times of switch signal overspeed are committed in total. The overall frequency of switch signal overspeed is $8.46\times10^{-4}$ times per kilometer.

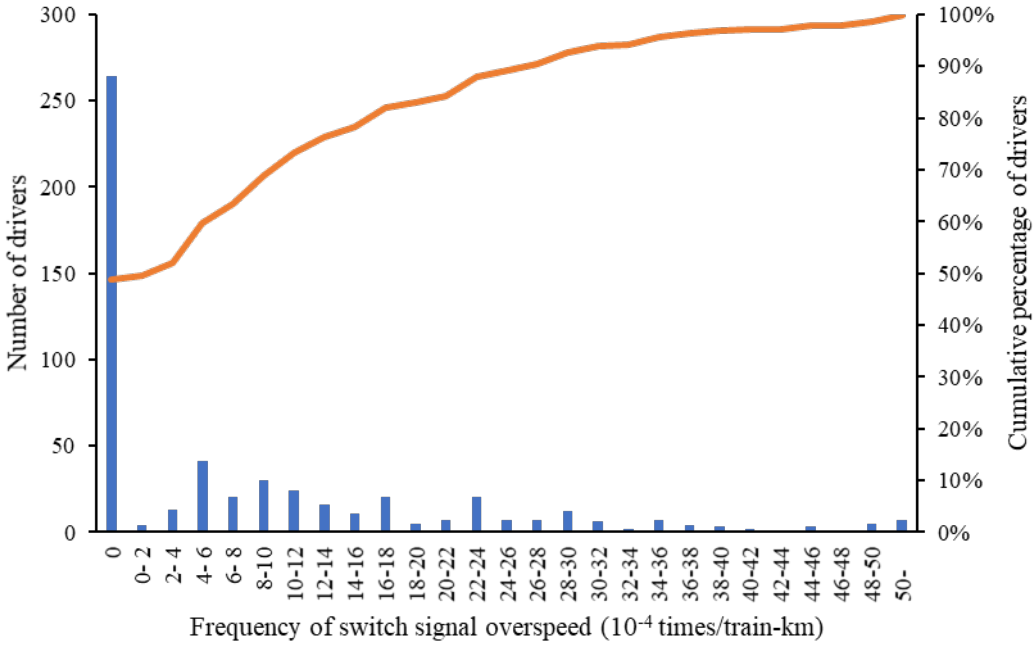


Figure 4-9: Histogram of the frequency of switch signal overspeed



doi:10.6342/NTU202001876

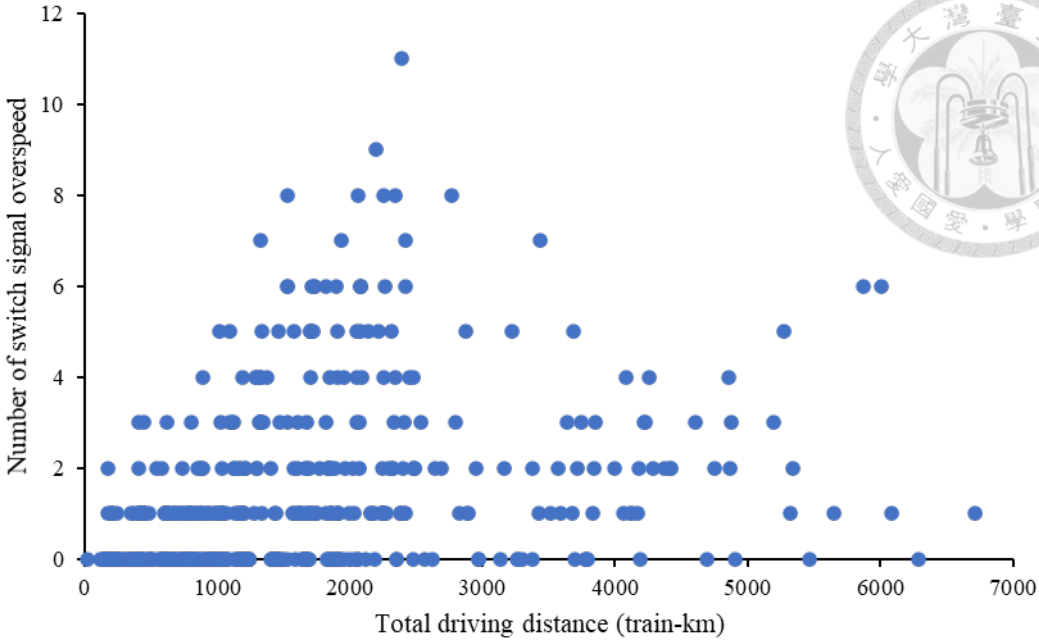


Figure 4-10: Scatter plot of approach signal overspeed and driving distance

Figure 4-9 illustrates the distribution of frequency of switch signal overspeed. The incidence of switch signal overspeed is discretized by $2.5\times10^{-4}$ times/train-km for all drivers. In this figure, 48.9% of drivers have no records on switch signal overspeed. Still, some outliers with much higher frequency of switch signal overspeed can be found in the figure. Figure 4-10 presents the relationship between switch signal overspeed and the driving distance. Moreover, this figure shows no significant trend for switch signal overspeed, and some extreme value might result from low total driving distance. According to the high percentage of non-identified data and low overall frequency, it could be derived that the data is not sufficient to find a trend line.



doi:10.6342/NTU202001876

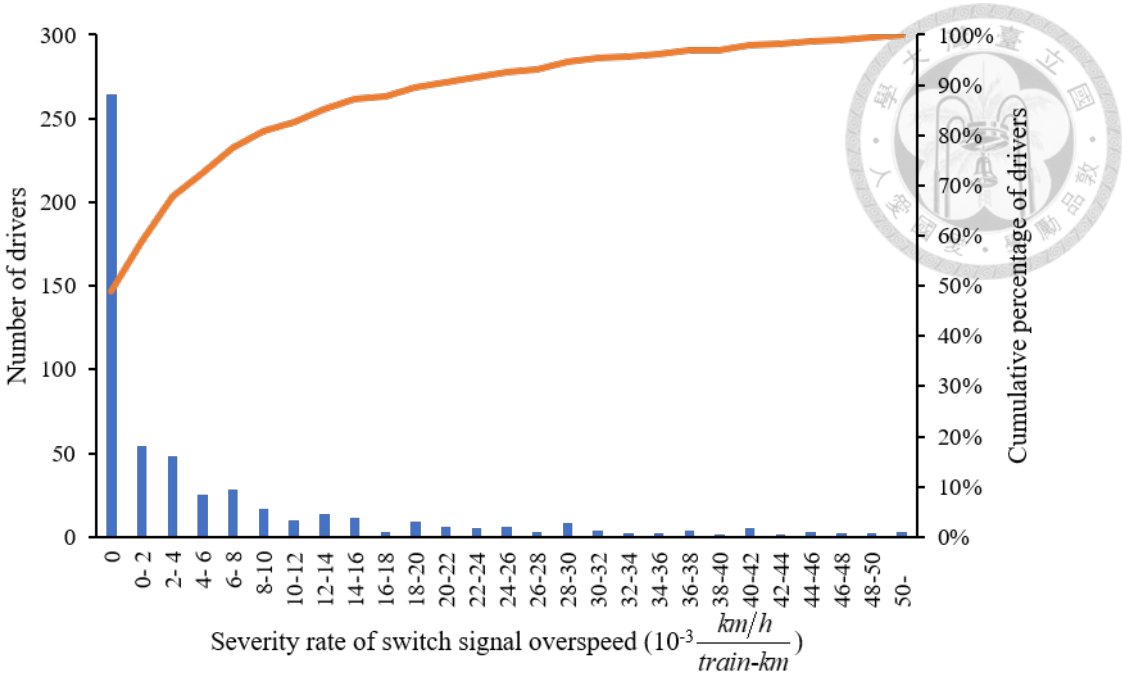


Figure 4-11: Histogram of the severity rate of switch signal overspeed

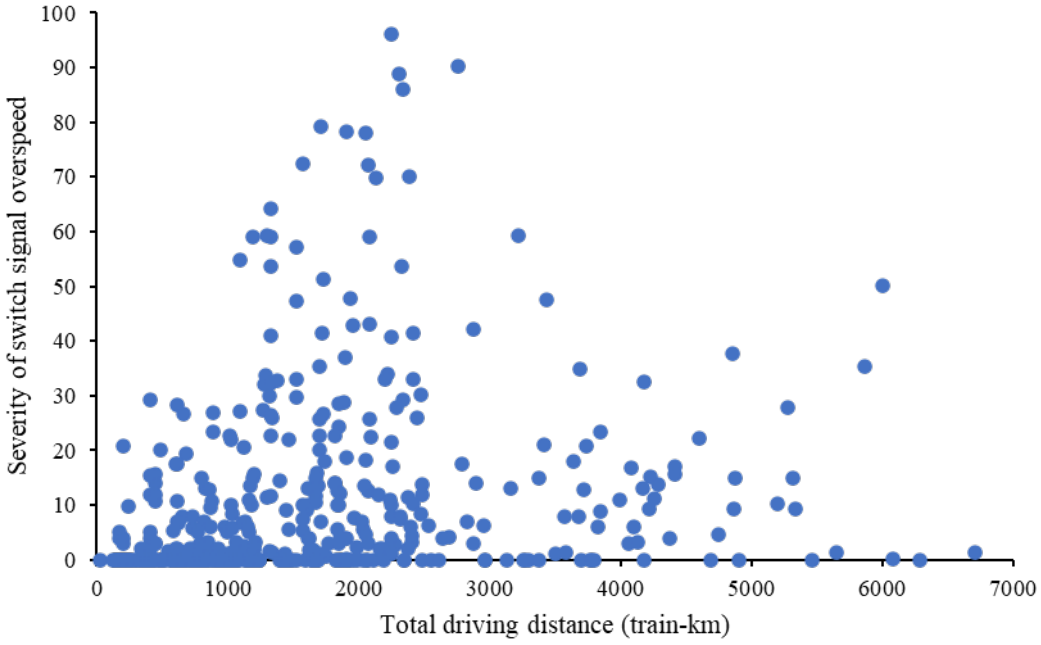


Figure 4-12: Scatter plot of severity of approach signal overspeed and driving distance

Figure 4-11 and Figure 4-12 show the result of severity of switch signal overspeed. According to Figure 4-10 and Figure 4-12, there is no significant difference of distribution between the two figures. Similar to approach signal overspeed, it can also be inferred that excessive train speed over the speed limit of switch signal does not



doi:10.6342/NTU202001876

varies significantly by drivers. For the normal drivers, excessive train speed over the speed limit of switch signal is quite similar. Only a few drivers who drives relying on ATP have much more severity of switch signal overspeed compared with the others.

### 4.2.4 Operational Overspeed

Among the 811,004.7 train-km operation distance of the 540 drivers, 128,083 seconds of operational overspeed are committed in total. The overall frequency of operation overspeed is 0.158 seconds per kilometer.

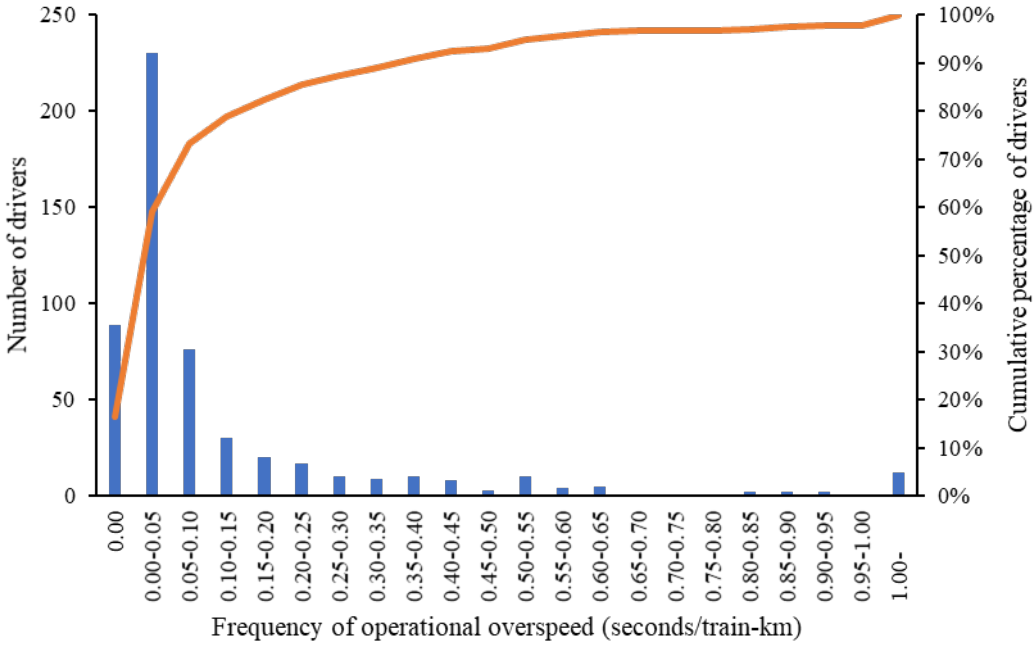


Figure 4-13: Histogram of the frequency of operational overspeed



doi:10.6342/NTU202001876

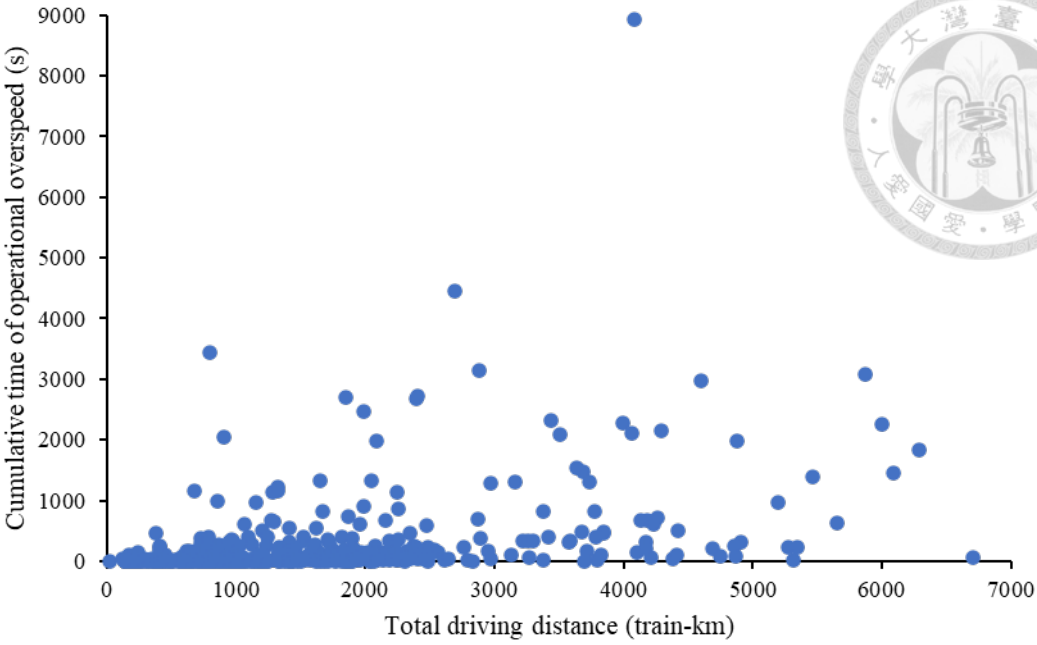


Figure 4-14: Scatter plot of operational overspeed and driving distance

Figure 4-13 depicts the distribution of frequency of operational overspeed. The incidence of operational overspeed is discretized by 0.05 seconds/train-km for all drivers. This figure presents that approximately 82% of drivers' frequency of operational overspeed is lower than 0.2 seconds/train-km. In contrast, much more outliers that overspeed far more frequently compared with ordinary drivers, and these outliers greatly affect the overall performance of operational overspeed. This result is also illustrated in Figure 4-14. The driving distance of these extreme drivers varies from 500 to 6000 train-km, and the one with almost twice as much time as the second-most overspeed driver had been driving over 4,000 train-km during the period. Such facts indicate that some drivers in TRA might be accustomed to overspeeding during operation.



doi:10.6342/NTU202001876

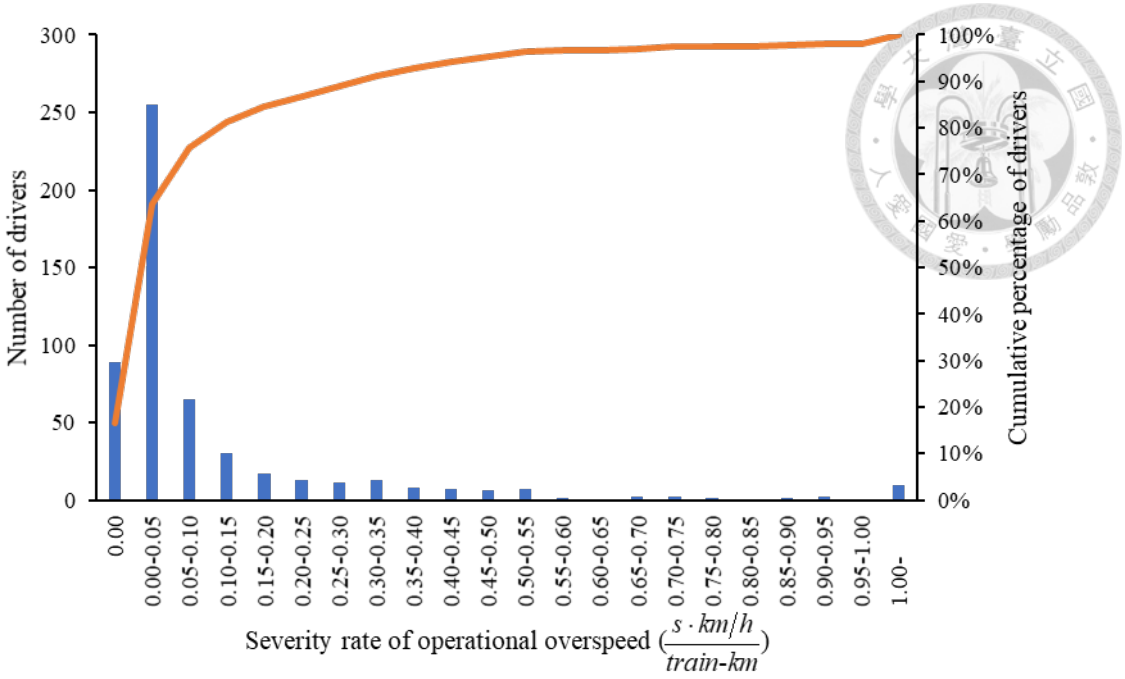


Figure 4-15: Histogram of the severity rate of operational overspeed

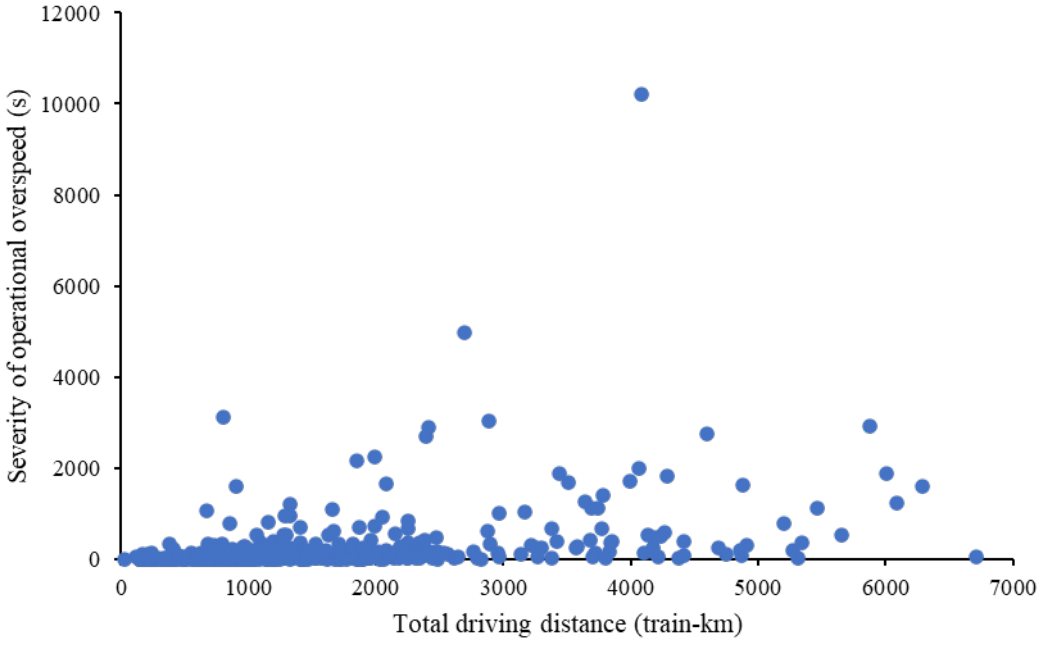


Figure 4-16: Scatter plot of severity of operational overspeed and driving distance

Figure 4-15 and Figure 4-16 show the result of severity of approach signal overspeed. According to Figure 4-14 and Figure 4-16, there is no significant difference of distribution between the two figures. It can be resulted from the operation logic of ATP. Under the protection of ATP, the train speed cannot exceed ATP permitted speed



doi:10.6342/NTU202001876

by 3 km/h, or the brake will be activated to slow down the train speed. Consequently, the excessive train speed over the ATP permitted speed is always less than 3 km/h. Thus, the distributions of the cumulative time of operational overspeed and the severity of operational overspeed are quite similar.

### 4.2.5 ATP Service Brake

Among the 811,004.7 train-km operation distance of the 540 drivers, 173 times of ATP service brake are committed in total. The overall frequency of ATP service brake is $2.13\times10^{-4}$ times per kilometer.

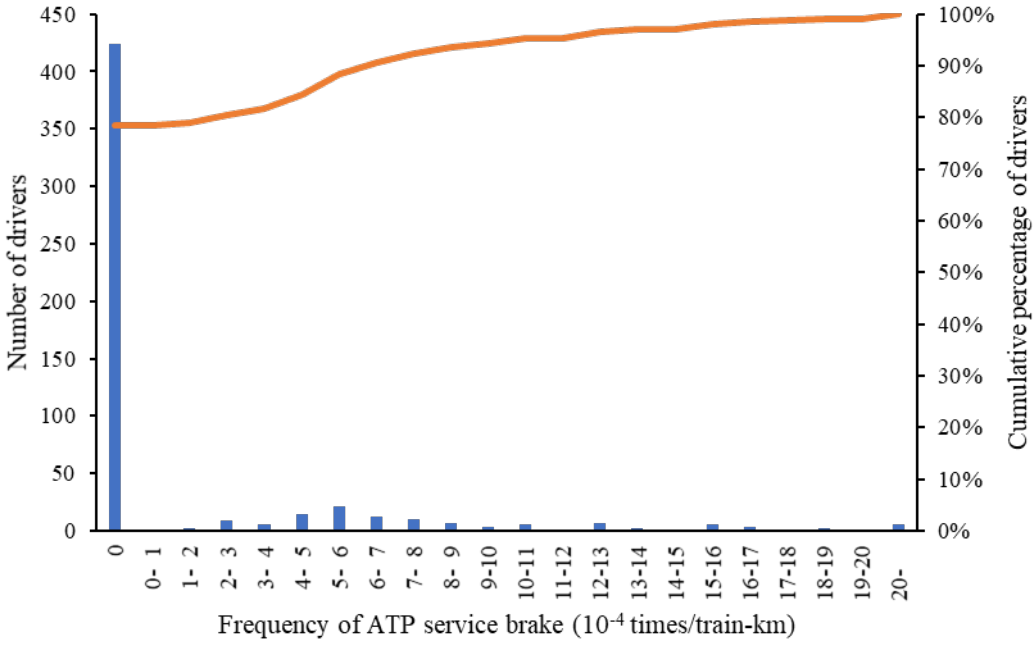


Figure 4-17: Histogram of the frequency of ATP service brake



doi:10.6342/NTU202001876

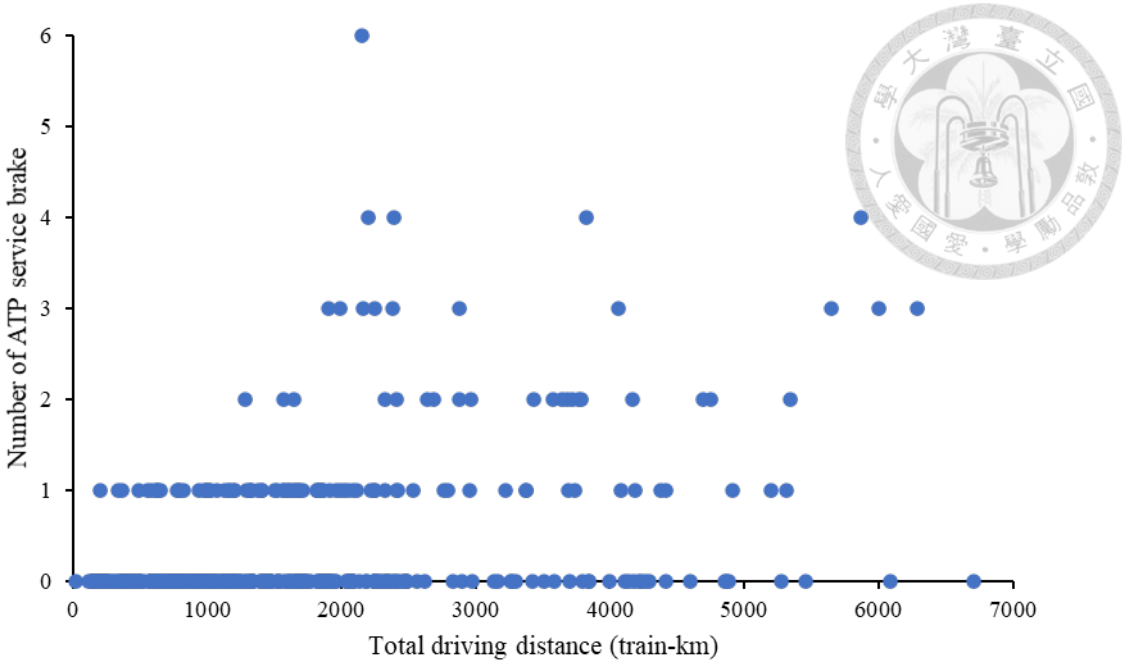


Figure 4-18: Scatter plot of ATP service brake and driving distance

Figure 4-17 presents the distribution of the frequency of ATP service brake. ATP service brake is discretized by $1\times10^{-4}$ times/train-km for all drivers. As shown in this figure, nearly 80% of drivers did not experience ATP service brake during this period. The trend in Figure 4-18 is not significant, either. Due to the low incidence and frequency of ATP service brake, it could be inferred that the sample size for ATP service brake is not enough for analysis, and it results in some extreme value in the data.

### 4.2.6 ATP Emergency Brake

Among the 811,004.7 train-km operation distance of 540 drivers, 37 times of ATP emergency brake are committed in total. The overall frequency of ATP service brake is $4.56\times10^{-5}$ times per train-km.



doi:10.6342/NTU202001876

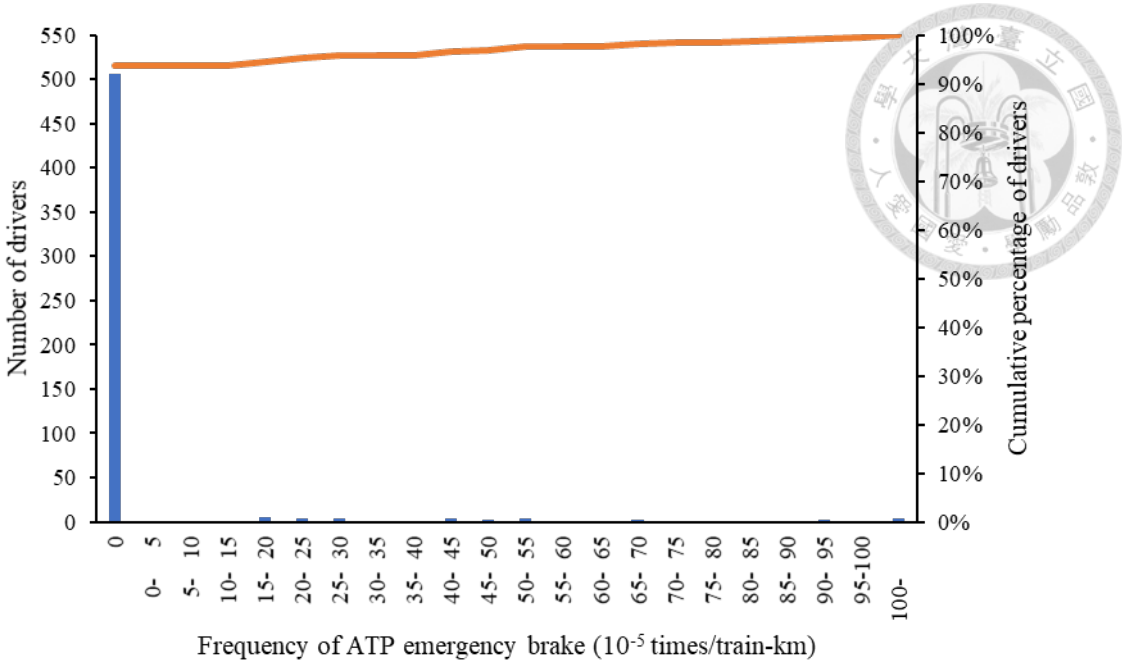


Figure 4-19: Histogram of the frequency of ATP emergency brake

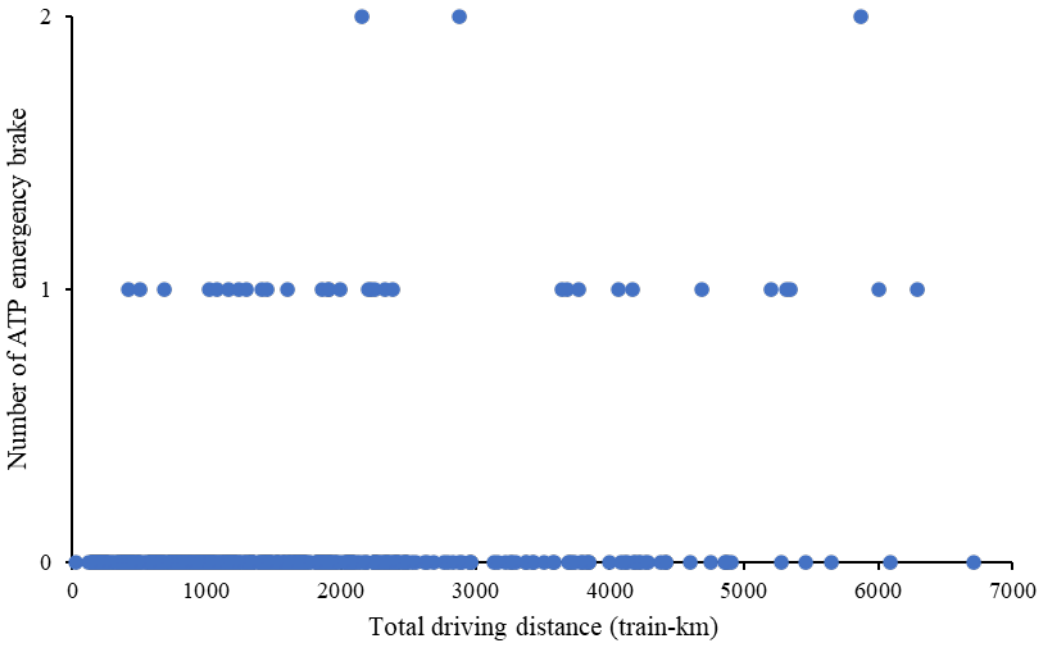


Figure 4-20: Scatter plot of ATP emergency brake and driving distance

Figure 4-19 presents the distribution of the frequency of ATP emergency brake. ATP emergency brake is discretized by $5\times10^{-5}$ times/train-km for all drivers. According to this figure, the characteristics of ATP emergency brake is identical to ATP emergency brake. Only 6% of driver had experienced ATP emergency brake during the period. The



doi:10.6342/NTU202001876

trend line in Figure 4-20 is not significant. According to the low incidence and frequency of ATP emergency brake, it could be inferred that the sample size for ATP emergency brake is not enough for analysis, and it results in some extreme value in the data.

### 4.2.7 IRIDB Evaluation

Table 4-2: Results of the AHP questionnaire

| High-risk driving behaviors | Weighting | Rank |
|---|---|---|
| Deceleration after target indication | 0.041 | 6 |
| Approach signal overspeed | 0.064 | 5 |
| Switch signal overspeed | 0.116 | 4 |
| Operational overspeed | 0.138 | 3 |
| ATP service brake | 0.202 | 2 |
| ATP emergency brake | 0.439 | 1 |

maximum eigenvalue = 6.144

C.R. = 0.023 < 0.1

An AHP designed questionnaire survey was distributed to 12 experts, including driving instructors at the TRA, and experts from the supervision agencies and research agencies. The result of AHP analysis is demonstrated in Table 4-2. The weighting of the ATP emergency and service brakes are substantially higher than the others among the six high-risk actions. Based on the operation logic of ATP, not until an emergency would ATP activate the brake to lower the train speed. Thus, these two behaviors have higher weightings as expected. The third and fourth most severe behaviors (i.e., operational overspeed and switch signal overspeed, respectively) have relatively similar value of weightings. The difference in severity may result from train speed. For switch signal



doi:10.6342/NTU202001876

overspeed, speed limit does not exceed 60 km/h. Hence, train speed is generally lower than operational overspeed, and severity may also be lower. Compared with switch signal overspeed, approach signal overspeed is relatively mild. For approach signals, the aim of speed limit is to ensure that the train can stop in front of the next stop signal, and run through it not necessarily causes accidents. For switch signals, the aim of speed limit is to ensure that the train can go through a turnout safely. Derailment is possible if a train commits switch signal overspeed. Lastly, the mildest high-risk driving behavior is deceleration after target indication. Although "deceleration after target indication" is considered an indicator of reliance on ATP, it does not actually result in any hazardous outcome. Therefore, deceleration after target indication has the lowest value of severity among the six high-risk driving behaviors.

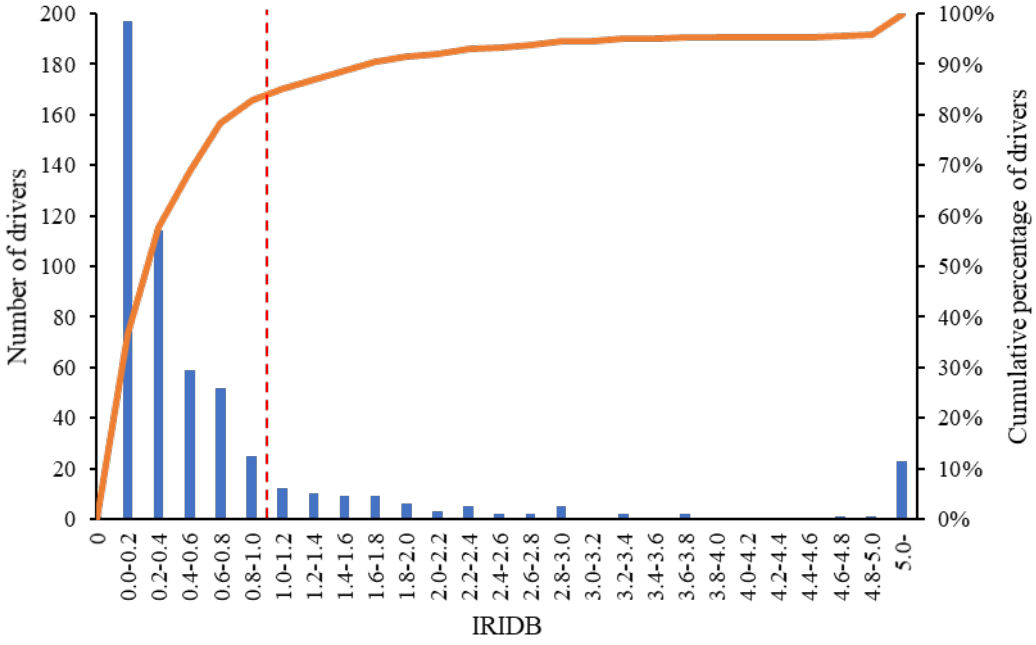


Figure 4-21: Histogram of IRIDB



doi:10.6342/NTU202001876

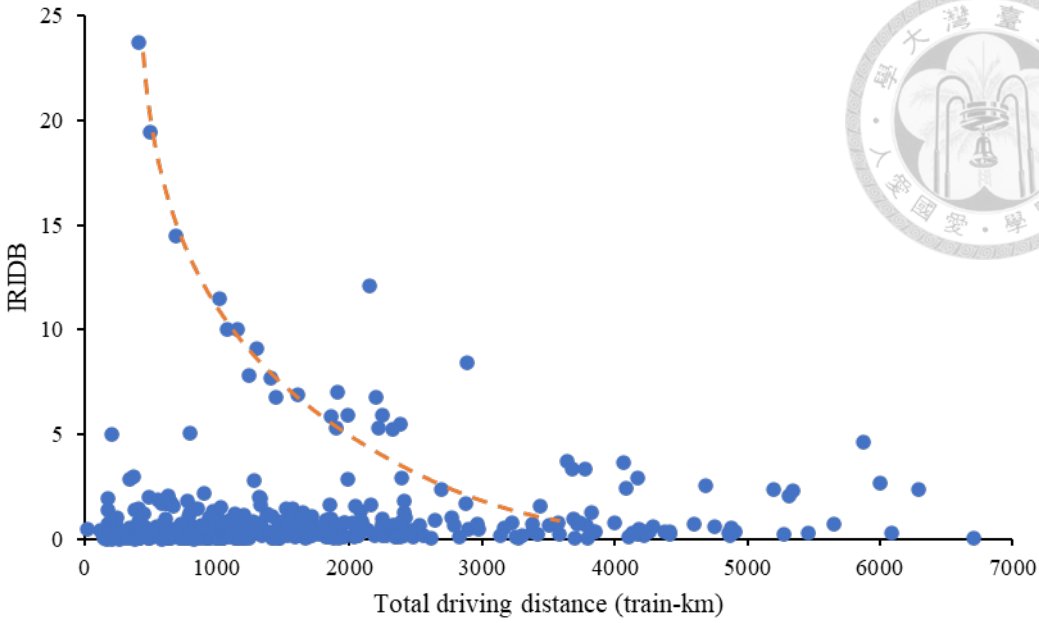


Figure 4-22: Scatter plot of IRIDB and driving distance

Figure 4-21 presents the performance in terms of the IRIDB. The red dash line denotes IRIDB = 1, representing average performance of all drivers. The distribution is similar to high-risk driving behaviors, over 80% if drivers whose IRIDB are less than 1. Nonetheless, there are still some drivers whose IRIDB are much higher than average. There are about 5% of drivers with IRIDB bigger than 5.

Figure 4-22 depicts the relationship between the IRIDB and the total driving distance. An inverse proportionality between the driving distance and the IRIDB could be observed for drivers with an IRIDB bigger than 5, as the orange line illustrated in Figure 4-22. Figure 4-23 indicates the high correlation between the IRIDB and ATP emergency brake. Such phenomenon might be resulted from the insufficient operation mileage for ATP emergency brake. As discussed in Section 4.2.6, due to the low incidence of ATP emergency brake, the indicator of ATP emergency brake varies drastically for the drivers with different operation mileage. Moreover, a relatively large weighting of ATP emergency brake lets the IRIDB easier to be affected by it. As a result, the IRIDB should be applied to driver's data with a longer driving distance.



doi:10.6342/NTU202001876

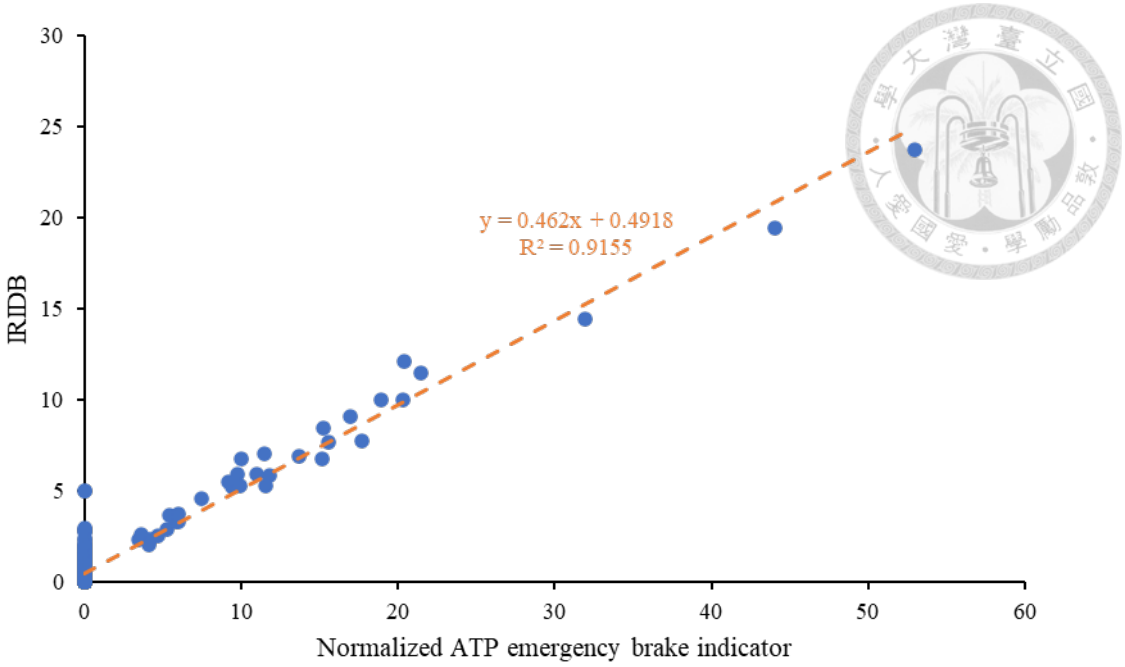


Figure 4-23: Relationship between normalized ATP emergency brake and risk index

### 4.2.8 Summary of High-Risk Driving Behavior Analysis

In the case study of high-risk driving behavior analysis, the high-risk driving behaviors committed by each driver is identified, and the indicators of the high-risk driving behaviors are evaluated. Moreover, the IRIDB combines the six indicators and provides a method to compare performance considering different high-risk driving behaviors between drivers, making it easier for operators to filter low-performance drivers and take countermeasures to improve safety.

Figure 4-24 and Figure 4-25 presents the risk of different types of high-risk driving behaviors with top 10% IRIDB, respectively. The red lines in these two figures depicts IRIDB = 1, representing average risk of all drivers. Figure 4-24 shows that the proportion of the ATP emergency brake is the largest. This may result from the insufficient data of ATP emergency brake. As was discussed in section 4.2.6, the frequency of ATP service brake is 4.56×10-5 times per train-km. Such low incidence results in highly imbalance of ATP emergency brake data. Moreover, according to the



doi:10.6342/NTU202001876

AHP survey, ATP emergency brake is the most severe high-risk behavior, and its weighting is more than twice as large as the second-most severe behavior. Consequently, Figure 4-25, depicts risk composition of drivers with top 10% highest risk index without ATP emergency brake to lower bias of it. With Figure 4-25, the operators can identify high-risk drivers and high-risk behaviors that the drivers commit frequently. And it can be applied to hold trainings for those high-risk drivers.

Note that IRIDB of the driver with 94th highest IRIDB (17th percentile) is below 1, representing that the risk of this driver is below average. The summation of IRIDB of drivers with the top 20% IRIDB accounts for 73.9% of the summation of every driver's IRIDB. Such a phenomenon fulfills the Pareto principle: 80% of the risk is controlled by 20% of the factors. This case study reveals that to improve the overall risk of driving behaviors, the minority drivers with high risk index should be targeted.



doi:10.6342/NTU202001876

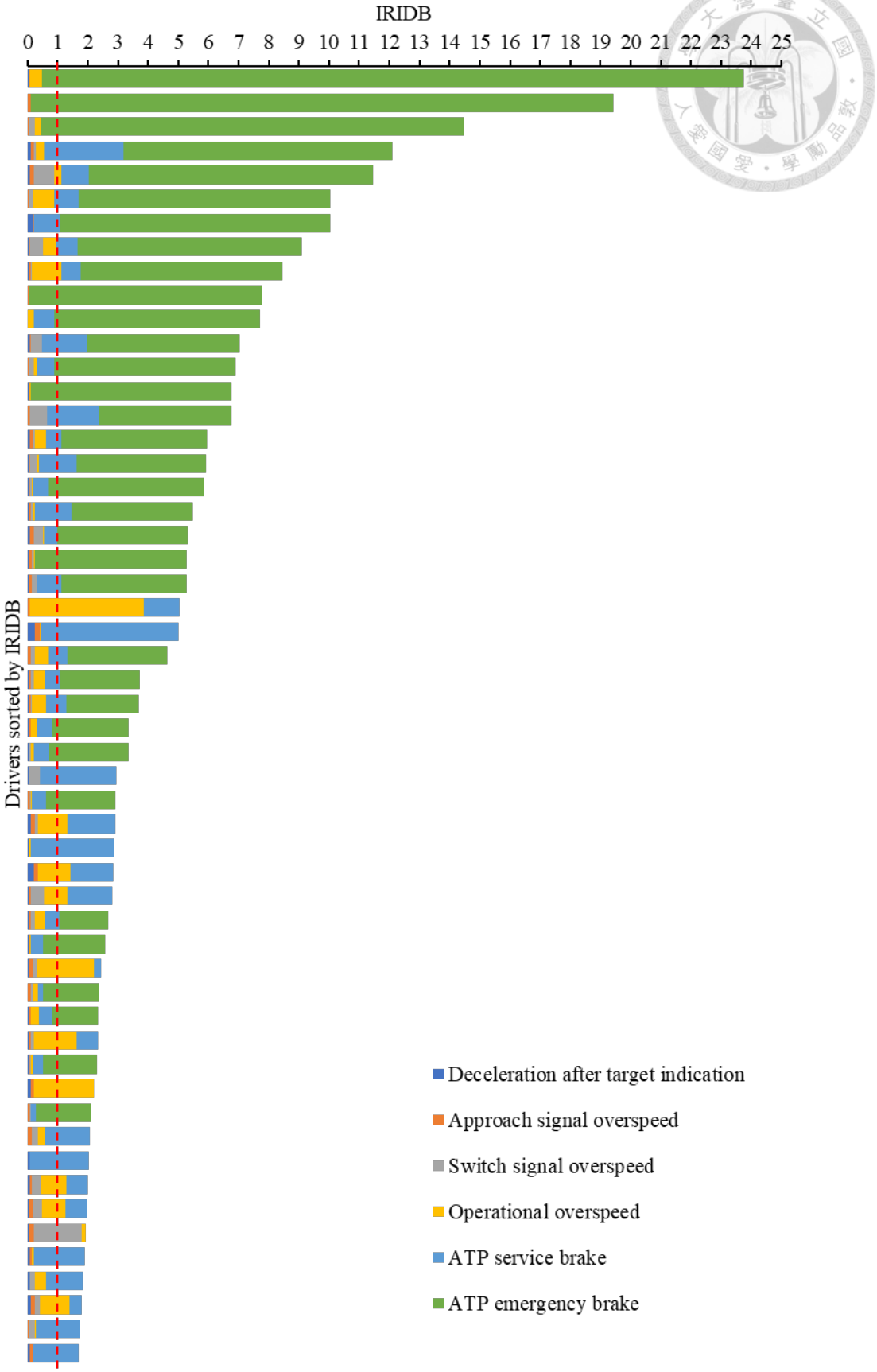


Figure 4-24: Risk composition of drivers with top 10% highest IRIDB



doi:10.6342/NTU202001876

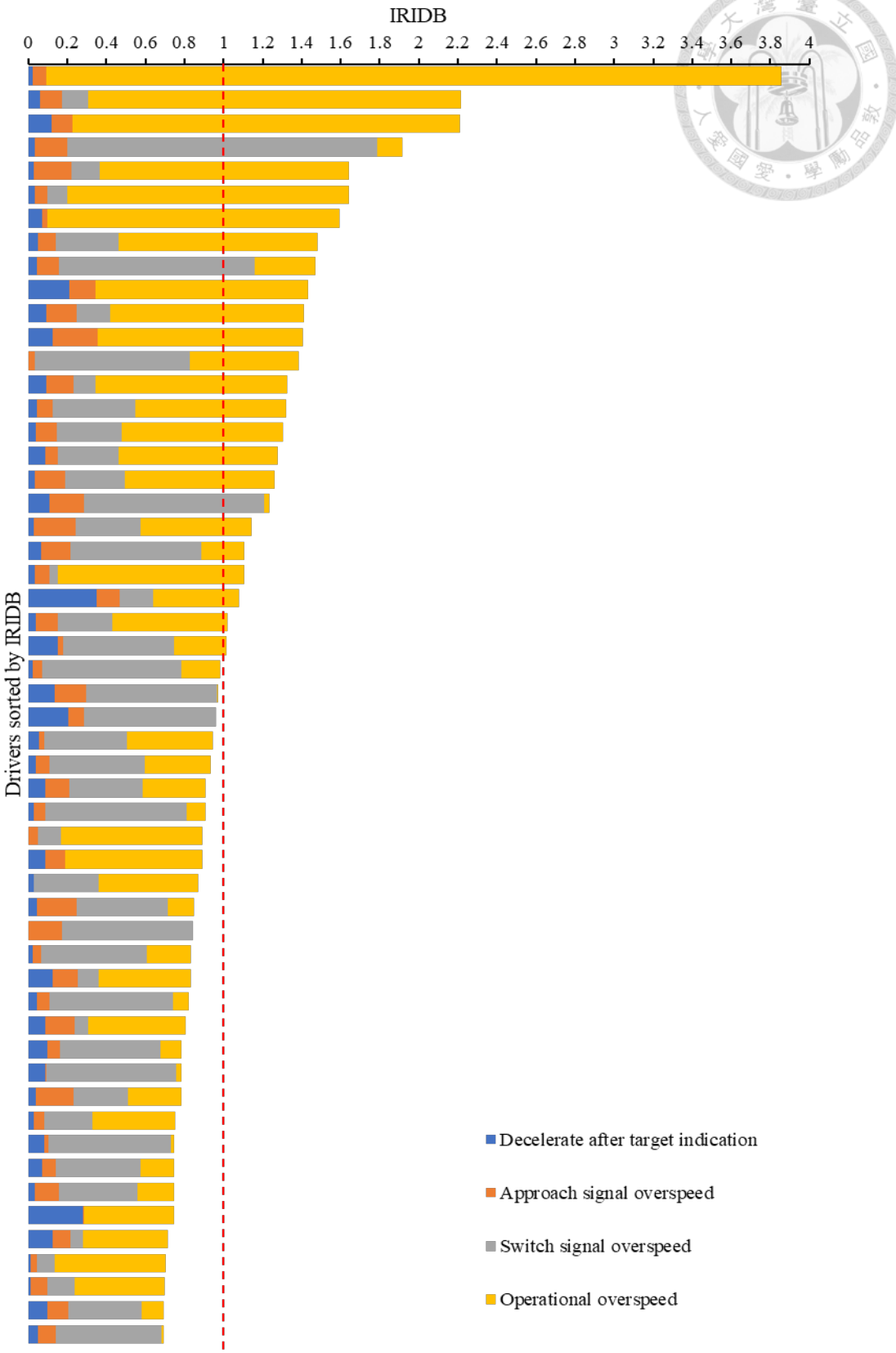


Figure 4-25: Risk composition of drivers with top 10% highest IRIDB

(without ATP emergency brake)



doi:10.6342/NTU202001876

## 4.3 Analysis of High-Risk Section

In the second case, the analysis of sections is performed to assess sections that high-risk driving behaviors occur frequently. Data used for the section analysis is ATP driving record figures of Puyuma express trains that passing through Yilan Line and North-link Line (Shulin to Hualien is used in this research) from May 2019 to July 2019. The data used in this case are data filtered from the same data used in section 4.2. The analysis section is 204.5 km in length, containing 142 blocks. The average length of blocks is 1.44 km, and standard deviation is 0.56 km. The length of blocks ranges from 0.61 km to 3.37 km. Figure 4-26 is the box plot of block length in section analysis. There are 999 effective data for section analysis, 827 of which are driving records between Shulin and Hualien, while the other 172 are the driving records between Qidu and Hualien.

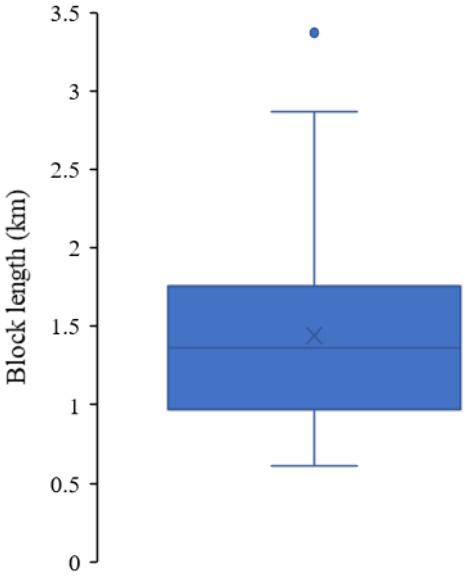


Figure 4-26: Box plot of block length between Shulin and Hualien



doi:10.6342/NTU202001876

### 4.3.1 Deceleration After Target Indication

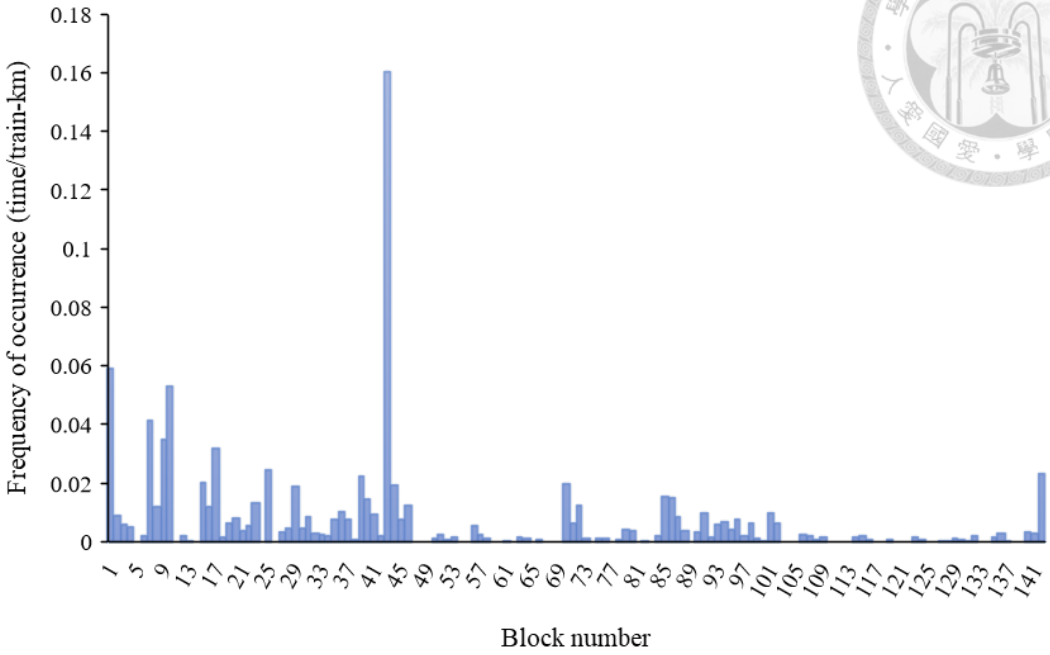


Figure 4-27: Histogram of frequency of deceleration after target indication by block

The result of deceleration after target indication identification by blocks is depicted in Figure 4-27. This figure reveals a section with much higher risk in deceleration after target indication, where the frequency is nearly three times as much as the section with the second-highest risk in deceleration after target indication. This section is located near Sandiaoling station, which is a section containing many curves with small radius. Moreover, sections with higher frequency of deceleration after target indication are concentrated in block 1 to block 45 and block 85 to block 105, located in Shulin to Shuangxi and Erjie to Dong'ao, respectively. Thus, it can be inferred that the frequency of deceleration after target indication might be correlated to curving sections.



doi:10.6342/NTU202001876

### 4.3.2 Approach Signal Overspeed

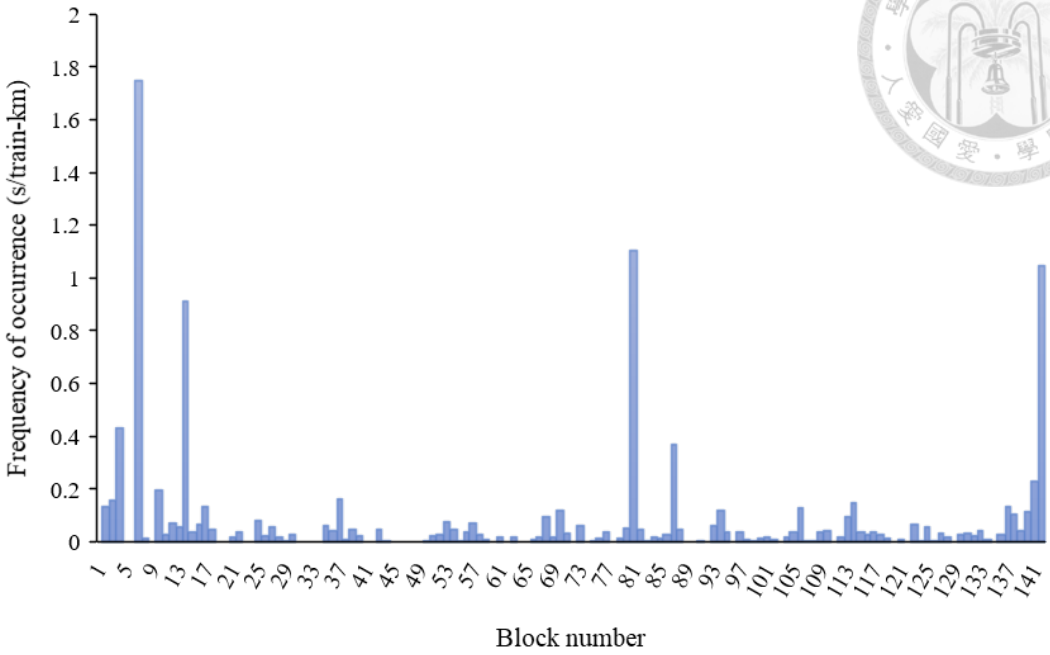


Figure 4-28: Histogram of frequency of approach signal overspeed by block

The result of approach signal overspeed identification by blocks is depicted in Figure 4-28. Blocks with a higher frequency of approach signal overspeed are concentrated in three sections: in block 1 to block 18, block 81 to 88, and block 136 to block 142, located in Shulin to Nangang, Yilan to Luodong and Jingmei to Hualien, respectively. These sections are close to stopping stations of Puyuma express trains, so that they may have higher chance to the approach signals in these sections.



doi:10.6342/NTU202001876

### 4.3.3 Switch signal Overspeed

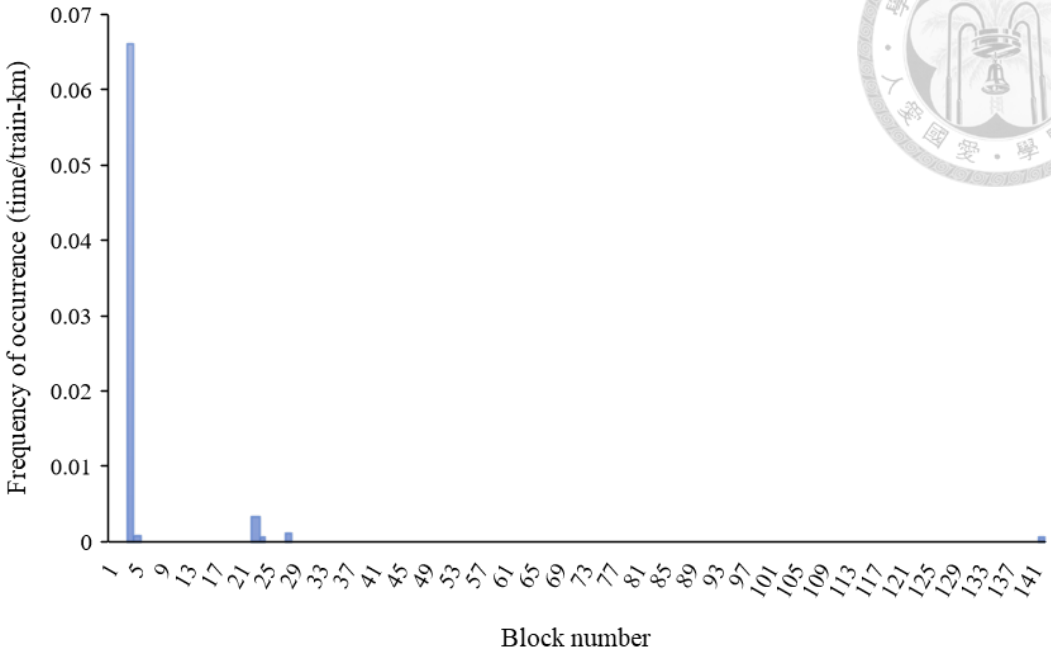


Figure 4-29: Histogram of frequency of switch signal overspeed by block

The result of switch signal overspeed identification by blocks is depicted in Figure 4-29. In this figure, switch signal overspeed only occurred in 6 blocks, and 53 out of 60 switch signal overspeed occurred in block 4 (near Banqiao station). Generally, switch signals are encountered when trains are going to stop at a station. Since the turnouts only locate in stations or junctions, switch signal overspeed only take place in these specific locations. Moreover, further analysis is necessary to figure out the reason to the abnormal number of switch signal overspeed that occur in the block. Due to the low occurrence of switch signal overspeed, more data is necessary for further analysis.



doi:10.6342/NTU202001876

### 4.3.4 Operational Overspeed

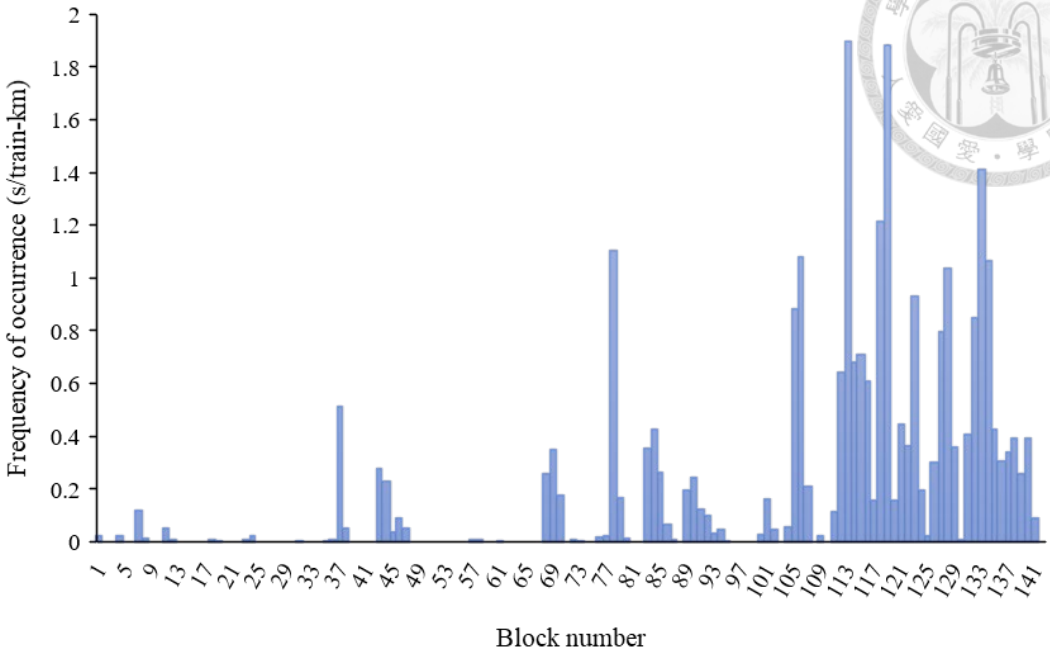


Figure 4-30: Histogram of frequency of operational overspeed by block

The result of operational overspeed identification by blocks is depicted in Figure 4-30. This figure presents the imbalance of sections that operational overspeed occurred. From block 1 to block 67, the frequency of operational overspeed is relatively low except in a few blocks. Such phenomenon may be resulted from two reasons. The first half of this section overlaps with West Main Line. The high frequency of trains results in a series of approach signals that slows down train speed. The second half of the section is the curving section of Yilan Line. Not only sharp curves make it difficult to operate too fast, the frequent change of speed limit also forces drivers to concentrate in order to control the train. On the other hand, frequency of operational overspeed rises drastically from block 68 to block 142. This section comprises plain section of Yilan Line and North-link Line. The relatively straight section with few stopping stations of Puyuma express trains makes drivers tend to operate trains in a faster speed, and the low control demand results in monotony so that drivers may be less cautious of overspeeding.



doi:10.6342/NTU202001876

### 4.3.5 ATP Service Brake

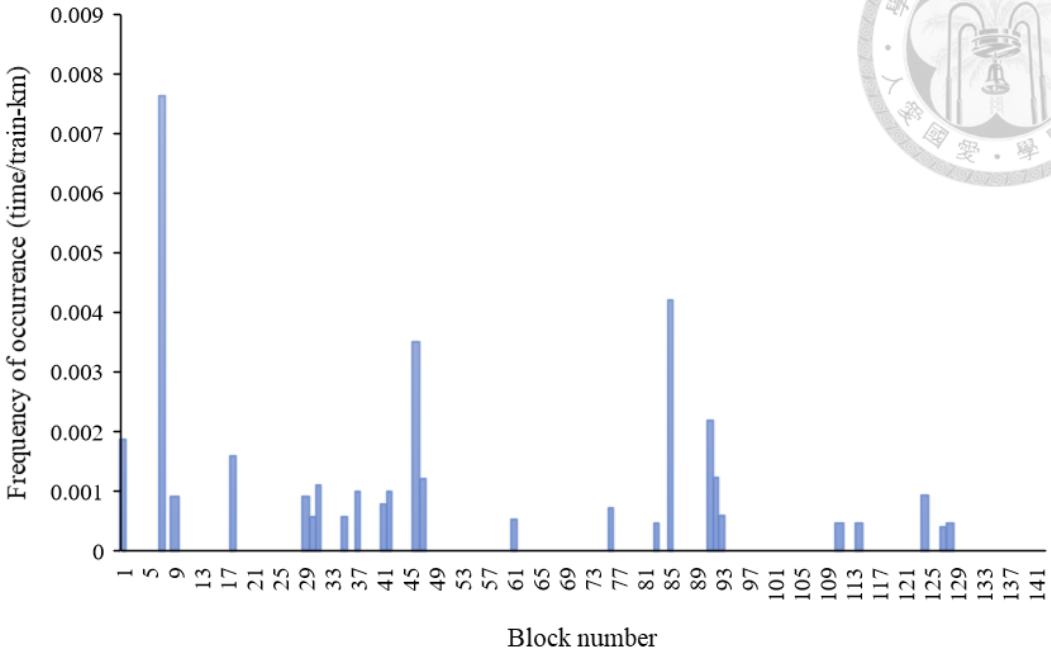


Figure 4-31: Histogram of frequency of ATP service brake by block

The result of ATP service brake identification by blocks is depicted in Figure 4-31. The blocks where ATP service brake occurred distribute evenly along the whole section. However, there are three blocks where ATP service brake occurred more frequently. Due to the low occurrence of ATP service brakes, more data is necessary for further analysis.



doi:10.6342/NTU202001876

### 4.3.6 ATP Emergency Brake

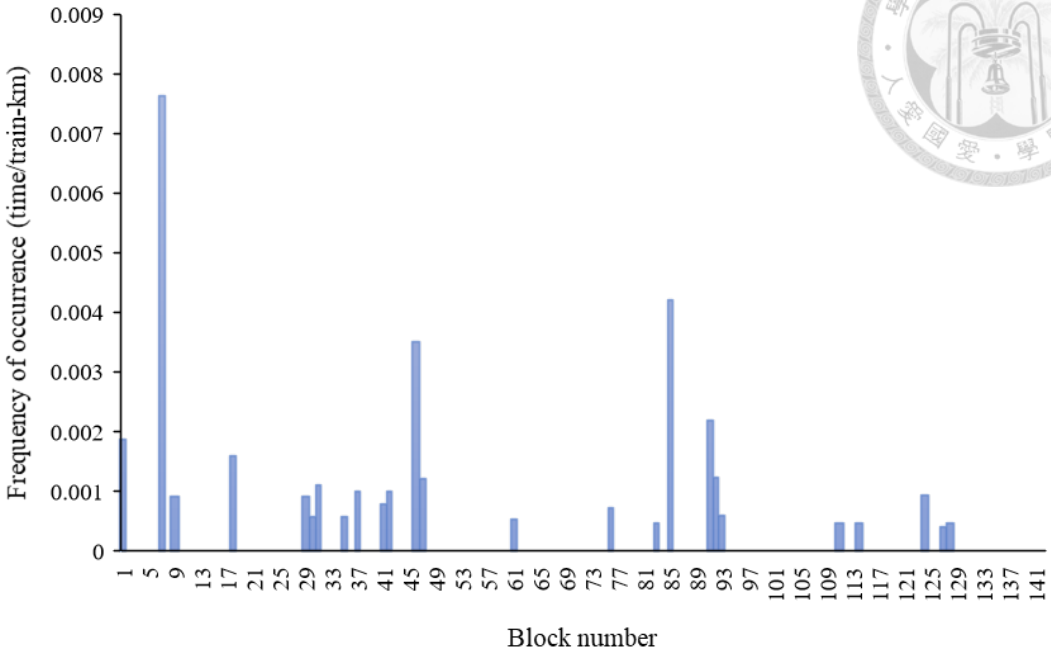


Figure 4-32: Histogram of frequency of ATP emergency brake by block

The result of ATP emergency brake identification by blocks is depicted in Figure 4-32. There are only 7 ATP emergency brakes in the analysis section in total. It is noted that 3 ATP emergency brake occurred in block 85, in which most ATP service brake take place. Due to the low occurrence of ATP emergency brakes, more data is necessary for further analysis.



doi:10.6342/NTU202001876

### 4.3.7 IRIDB Evaluation

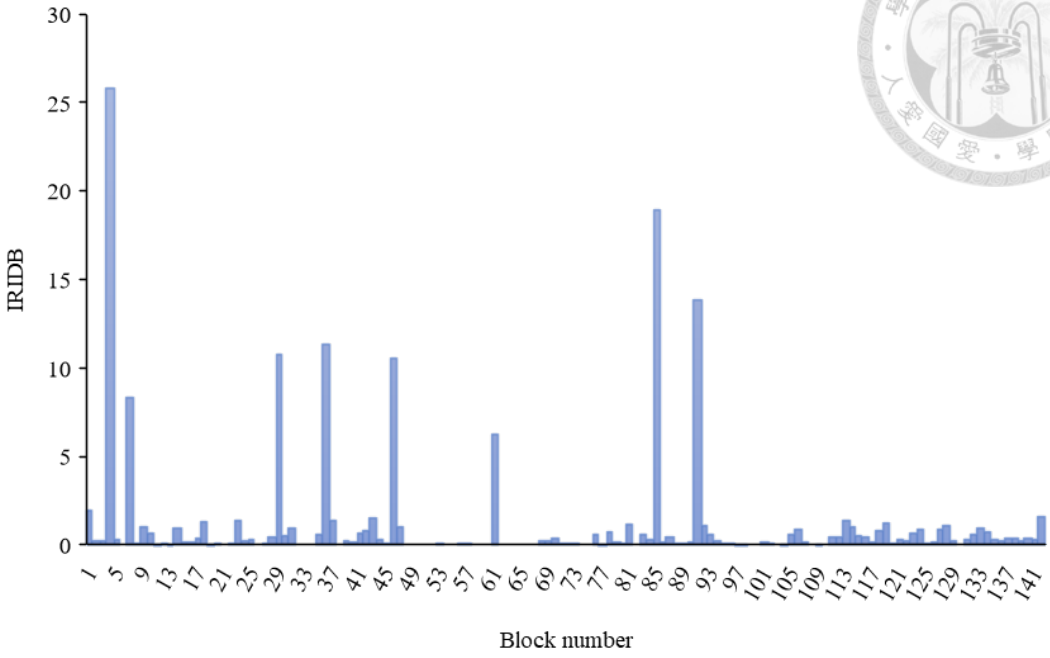


Figure 4-33: Histogram of IRIDB by blocks

The result of IRIDB by blocks is demonstrated in Figure 4-33. Extreme values of high-risk driving behaviors revealed in Sections 4.3.1 to 4.3.6 are reflected in the figure. Figure 4-34 presents the distribution of the IRIDB by blocks. This figure shows that risk indices of a majority of blocks are very low. The IRIDB of 54 out of 142 blocks (38.0%) are less than 0.1, and the IRIDB of 99 out of 142 blocks (69.7%) are less than 0.5. Only 25 out of 142 blocks (17.6%) have an IRIDB is higher than 1.



doi:10.6342/NTU202001876

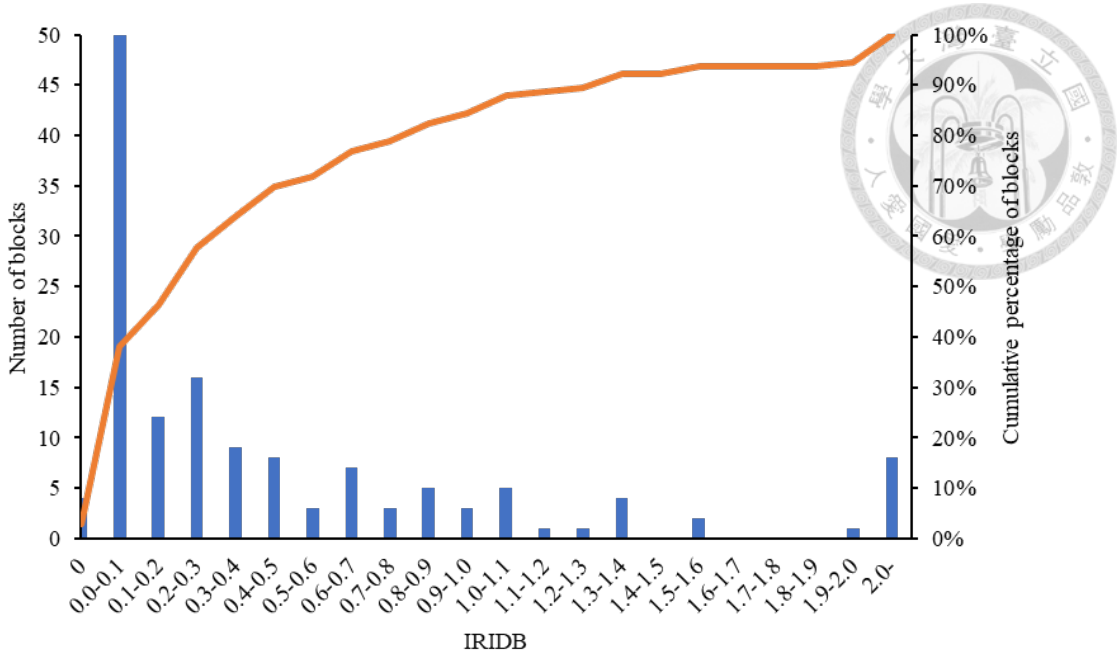


Figure 4-34: Distribution of IRIDB by blocks

### 4.3.8 Summary of High-Risk Section Analysis

In the case study of high-risk section analysis, locations that high-risk driving behaviors are often committed are identified, and the indicators of them are evaluated. Moreover, the IRIDB, which combines the six indicators, provides a method to compare risk considering different high-risk driving behaviors between different sections, making it easier for operators to find out sections with higher risk and take countermeasures to improve safety.

Figure 4-35 shows the risk composition of every block along the section between Shulin and Hualien. Figure 4-36 illustrates the risk composition of the top 10 sections with the highest IRIDB. The red lines in these two figures depicts IRIDB = 1, representing the average risk of all blocks. Similar to the analysis of high-risk driving behavior, ATP emergency brake and ATP service brake account for the majority proportion of the sections with high IRIDB. Another similarity to the analysis of high-risk driving behavior is the extreme distribution of IRIDB for all blocks. Only 22 out



doi:10.6342/NTU202001876

of the 142 blocks (15.9%) have risk above average.

The summation of IRIDB of the sections with the top 20% IRIDB accounts for 84.3% of summation of every section's IRIDB, which also fulfills the Pareto principle. It can be summarized that the overall risk of the entire network is controlled by some small but critical sections in this case study. Safety improvements to these sections are the most efficient means to enhance the overall safety for the network.



doi:10.6342/NTU202001876

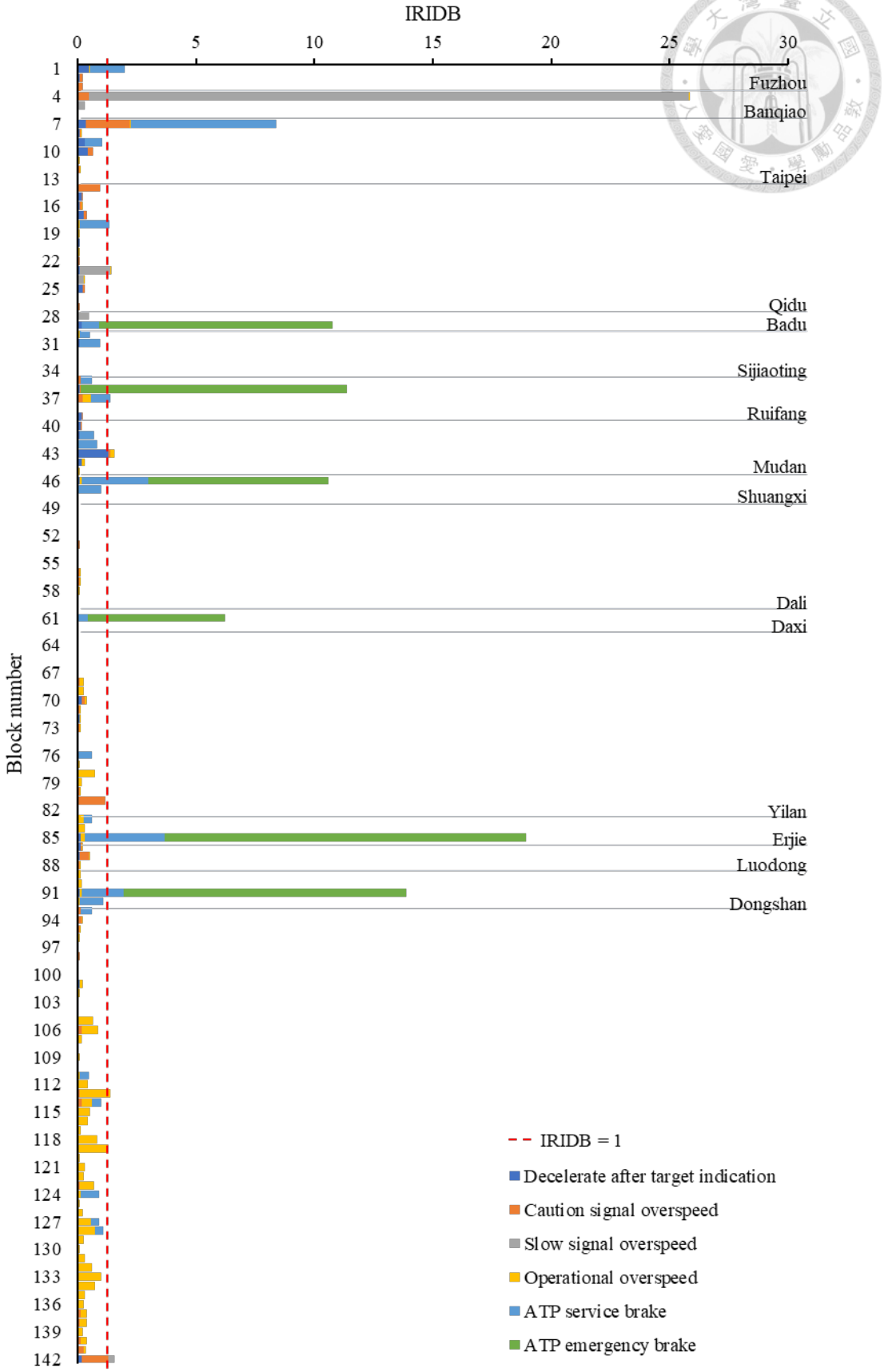


Figure 4-35: Risk composition of blocks



doi:10.6342/NTU202001876

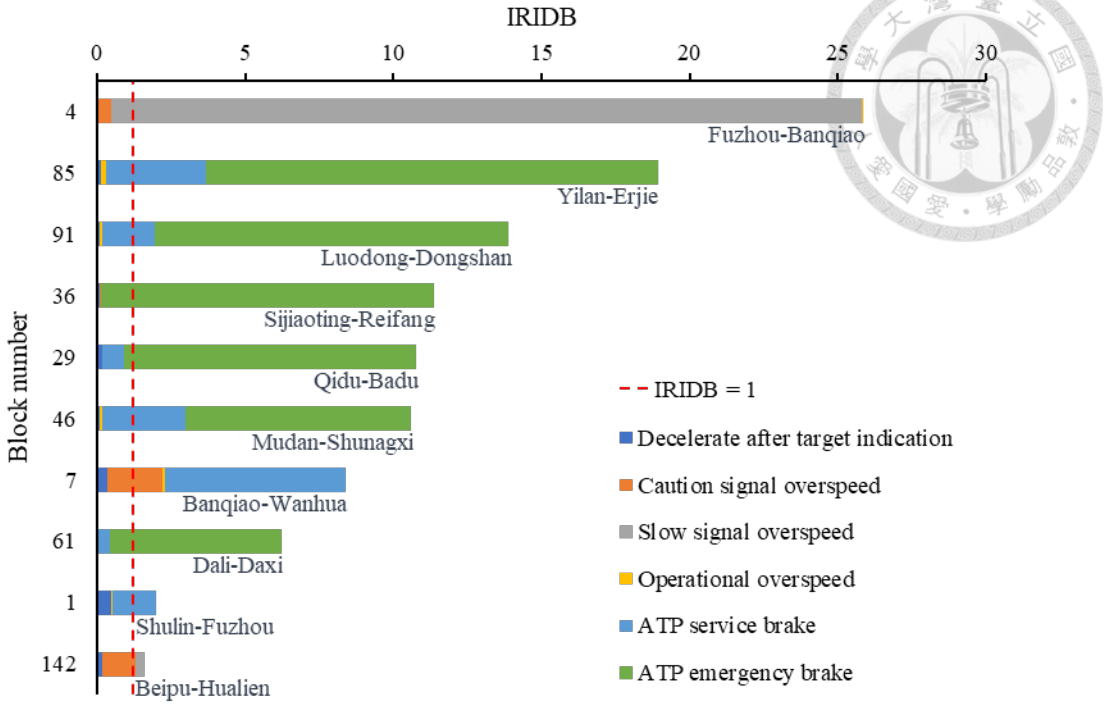


Figure 4-36: Risk composition of top 10 highest IRIDB blocks

## 4.4 Summary and Discussion

In this chapter, the two case studies demonstrate that the modules proposed in this research can not only identify the high-risk driving behaviors and compute the indicators thereof but also provide an integrated risk index for evaluating the driver and section safety. The results show that this proposed framework can assist railway operators identifying the high-risk drivers and sections. The integrated risk index for driving behaviors (IRIDB) is also computed to evaluate the overall risk for drivers and sections, so that antidotes can be taken for safety improvement on those with low performance. Moreover, the IRIDB in both cases exhibit the Pareto principle of risk in railway operation. For both drivers and sections, the distributions of IRIDB are extremely imbalanced. The majority of the drivers and sections are in a low level of IRIDB, while the minority of the drivers and sections has an IRIDB much higher than average. The result confirms the Pareto principle in high-risk driving behaviors, that the top 20% of the drivers and sections with higher IRIDB contains 80% of the risk.



doi:10.6342/NTU202001876

The phenomenon reveals that the 20% of drivers and sections with a higher IRIDB are important target of safety improvement. If the operators aim to improve the overall safety of the railway system, dealing with the 20% of drivers and sections with a higher IRIDB is the most efficient approach because the maximum achievement can be obtained with the minimum effort.



doi:10.6342/NTU202001876

# CHAPTER 5 CONCLUSION AND FUTURE WORK

With the development of railway technology, people gradually emphasize on safety of railway transportation. Despite the continuous improvement of safety facilities , no facility can guarantee 100% safety. In order to improve safety of railway systems other than safety facilities, this research turns to assess human factors during train operation. With the development of big data analysis, data collected during operation is widely used in different aspects. This research applies the existing data from ATP that records information during operation to assess the driver performance and section safety. Moreover, the IRIDB proposed in this research summarizes the different actions detected and sorts the drivers and sections by risk. Such index can help the operators to arrange the priorities of countermeasures.

## 5.1 Conclusion

This research proposes a framework of driver high-risk driving behavior detection based on the driving record data. The conclusions and contributions are summarized as follows:

(1) To demonstrate potential application of ATP record data

The ATP system (or ETCS Level 1 equivalent signaling system) has been developed and widely used in a variety of countries for over a decade. The system collects and saves much valuable data but few researchers or operators had tried to utilize it. This research provides a sample of ATP record data application, that assesses driver performance with the data.

(2) To provide a fully automatic process for driver behavior assessment

As mentioned in Section 1.1, several driving instructors are responsible for oversight of driver behavior. Despite the introduction of the new overspeed detection system for driver assessment, the functions of the new system are limited.



doi:10.6342/NTU202001876

As a result, this research establishes an automatic process to detect high-risk driving behaviors, which is more accurate and efficient.

(3) To assess sections with higher frequency of high-risk driving behaviors

Apart from assessing the driving performance for each driver, this research also assesses the driver performance in each section. The result can indicate sections where drivers tend to commit high-risk driving behaviors, so that the operators can take counterparts to improve section safety.

(4) To develop a model for driver and section risk evaluation

This research also constructs an AHP model to evaluate the IRIDB of the six high-risk driving behaviors. With this index, it becomes possible to arrange priorities of improvement works by severity.

## 5.2 Future Work

This section summarizes some applications that could be extended in the future. Although the developed process can identify high-risk driving behaviors automatically, there are still some restrictions and limitations to be improved. These future works are listed below.

(1) Access additional data to obtain a more objective IRIDB

As mentioned in Section 4.2.6, insufficient data of some high-risk driving behaviors with low incidence might result in bias while computing the frequency. A more objective IRIDB could be obtained with accessing more data.

(2) Investigate additional factors related to frequency of high-risk driving behaviors for drivers and sections

In this research, incidence of high-risk driving behaviors for drivers and sections is evaluated. However, further discussions such as the relationship between frequency and other probable factors are not performed in this research. To further



doi:10.6342/NTU202001876

improve safety, more detailed studies should be done. For example, futures studies may try to find out factors that may affect driving risk, to analyze the relationship between the drivers, the sections and the corresponding characteristics. With such analysis, the operators could focus on factors that are highly related to the risks, such as age, experience, slope, curvature, to find the right antidote. Apart from characteristics of drivers and sections, effects of delay conditions, different time, and some other factors are also worth researching in future studies.

(3) Develop an identification model for more high-risk driving behaviors with more detailed data

The ATP driving record data is applied for analyzing the high-risk driving behaviors in this research. Nevertheless, there are some other database that collects and stores driving record. For example, some trains in TRA are installed with the Train Control and Monitor System (TCMS). TCMS not only records train speed but also records the speed command, the brake command and other data. With this additional data, it is possible to improve the accuracy of the high-risk driving behavior identification and to discover more high-risk driving behaviors. Moreover, with the additional command data, it becomes possible to analyze the interactions between the drivers, the signals, and ATP.

(4) Develop a more objective model for risk index evaluation

In this research, an AHP structure is used to evaluate the weightings of risk index. However, the concept of AHP is still based on subjective opinion of experts. More detailed quantified risk analysis methods can be applied for the integrated risk index evaluation.



doi:10.6342/NTU202001876

## REFERENCE


Aven, T. (2008). A semi-quantitative approach to risk analysis, as an alternative to QRAs. *Reliability Engineering & System Safety, 93*(6), 790-797. doi:10.1016/j.ress.2007.03.025

Baysari, M. T., McIntosh, A. S., & Wilson, J. R. (2008). Understanding the human factors contribution to railway accidents and incidents in Australia. *Accident Analysis & Prevention, 40*(5), 1750-1757. doi:10.1016/j.aap.2008.06.013

Boehm, B. (1989). *Software risk management.* Paper presented at the ESEC '89, Berlin, Heidelberg.

Dorrian, J., Baulk, S. D., & Dawson, D. (2011). Work hours, workload, sleep and fatigue in Australian Rail Industry employees. *Applied ergonomics, 42*(2), 202-209. doi:10.1016/j.apergo.2010.06.009

Dorrian, J., Roach, G. D., Fletcher, A., & Dawson, D. (2007). Simulated train driving: fatigue, self-awareness and cognitive disengagement. *Applied ergonomics, 38*(2), 155-166. doi:10.1016/j.apergo.2006.03.006

El Rashidy, R., Hughes, P., Esteban, M. F., & Van Gulijk, C. (2018). *Automated train driver competency performance indicators using real train driving data.* Paper presented at the Annual European Safety and Reliability Conference.

European Committee for Electrotechnical Standardization (CENELEC). (1999). BS EN 50126:1999 Railway applications — The specification and demonstration of Reliability, Availability, Maintainability and Safety (RAMS) —. In. London: BSI.




doi:10.6342/NTU202001876

Figueres-Esteban, M., Hughes, P., & Van Gulijk, C. (2016). Visual analytics for text-based railway incident reports. *Safety science, 89*, 72-76. doi:10.1016/j.ssci.2016.05.009

Forman, E. H. (1990). Random indices for incomplete pairwise comparison matrices. *European Journal of Operational Research, 48*(1), 153-155. doi:10.1016/0377-2217(90)90072-J

Gnoni, M. G., & Lettera, G. (2012). Near-miss management systems: A methodological comparison. *Journal of Loss Prevention in the Process Industries, 25*(3), 609-616. doi:10.1016/j.jlp.2012.01.005

Guo, M., Wei, W., Liao, G., & Chu, F. (2016). The impact of personality on driving safety among Chinese high-speed railway drivers. *Accident Analysis & Prevention, 92*, 9-14. doi:10.1016/j.aap.2016.03.014

Health and Safety Executive. (2004). *Investigating accidents and incidents. A workbook for employers, unions, safety representatives and safety professionals*. (HSG 245). Retrieved from https://www.hse.gov.uk/pubns/hsg245.pdf

Hickey, A. R., & Collins, M. D. (2017). Disinhibition and train driver performance. *Safety science, 95*, 104-115. doi:10.1016/j.ssci.2017.02.016

Independent Transport Safety and Reliability Regulator (ITSRR). (2005). *Information paper: driver safety systems and automatic train protection*. [Sydney, NSW]: ITSRR.



doi:10.6342/NTU202001876

International Organization for Standardization (ISO). (2010). IEC/ISO 31010:2009 Risk management – risk assessment techniques. In. Brussels: CENELEC.

Jay, S. M., Dawson, D., Ferguson, S. A., & Lamond, N. (2008). Driver fatigue during extended rail operations. *Applied ergonomics, 39*(5), 623-629. doi:10.1016/j.apergo.2008.01.011

Jong, J.-C., Lai, Y.-C., Young, C.-C., & Chen, Y.-F. (2020). Application of Fault Tree Analysis and Swiss Cheese Model to the Overspeed Derailment of Puyuma Train in Yilan, Taiwan. *Transportation Research Record: Journal of the Transportation Research Board, 2674*(5), 33-46. doi:10.1177/0361198120914887

Kaeeni, S., Khalilian, M., & Mohammadzadeh, J. (2018). Derailment accident risk assessment based on ensemble classification method. *Safety science, 110*, 3-10. doi:10.1016/j.ssci.2017.11.006

Kaplan, S. (1990). On the inclusion of precursor and near miss events in quantitative risk assessments: A Bayesian point of view and a space shuttle example. *Reliability Engineering & System Safety, 27*(1), 103-115. doi:10.1016/0951-8320(90)90034-K

Kyriakidis, M., Majumdar, A., & Ochieng, W. Y. (2015). Data based framework to identify the most significant performance shaping factors in railway operations. *Safety science, 78*, 60-76. doi:10.1016/j.ssci.2015.04.010

Leitner, B. (2017). A General Model for Railway Systems Risk Assessment with the Use of Railway Accident Scenarios Analysis. *Procedia Engineering, 187*, 150-159. doi:10.1016/j.proeng.2017.04.361



doi:10.6342/NTU202001876

Macciotta, R., Martin, C. D., Morgenstern, N. R., & Cruden, D. M. (2016). Quantitative risk assessment of slope hazards along a section of railway in the Canadian Cordillera—a methodology considering the uncertainty in the results. *Landslides, 13*(1), 115-127. doi:10.1007/s10346-014-0551-4

MacKenzie, C. A. (2014). Summarizing Risk Using Risk Measures and Risk Indices. *Risk Analysis, 34*(12), 2143-2162. doi:10.1111/risa.12220

McKinnon, R. C. (2012). *Safety Management*. Boca Raton: CRC Press.

McLeod, R. W., Walker, G. H., & Moray, N. (2005). Analysing and modelling train driver performance. *Applied ergonomics, 36*(6), 671-680. doi:10.1016/j.apergo.2005.05.006

Mirabadi, A., & Sharifian, S. (2010). Application of association rules in Iranian Railways (RAI) accident data analysis. *Safety science, 48*(10), 1427-1435. doi:10.1016/j.ssci.2010.06.006

Myrtek, M., Deutschmann-Janicke, E., Strohmaier, H., Zimmermann, W., Lawerenz, S., Brügner, G., & Müller, W. (1994). Physical, mental, emotional, and subjective workload components in train drivers. *Ergonomics, 37*(7), 1195-1203. doi:10.1080/00140139408964897

Naweed, A. (2013). Psychological factors for driver distraction and inattention in the Australian and New Zealand rail industry. *Accident Analysis & Prevention, 60*, 193-204. doi:10.1016/j.aap.2013.08.022

Naweed, A. (2014). Investigations into the skills of modern and traditional train driving. *Applied ergonomics, 45*(3), 462-470. doi:10.1016/j.apergo.2013.06.006



doi:10.6342/NTU202001876

Ohlson, J. A., & Kim, S. (2015). Linear valuation without OLS: the Theil-Sen estimation approach. *Review of Accounting Studies, 20*(1), 395-435. doi:10.1007/s11142-014-9300-0

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Dubourg, V. J. t. J. o. m. L. r. (2011). Scikit-learn: Machine learning in Python. *12*, 2825-2830.

Punzet, L., Pignata, S., & Rose, J. (2018). Error types and potential mitigation strategies in Signal Passed at Danger (SPAD) events in an Australian rail organisation. *Safety science, 110*, 89-99. doi:10.1016/j.ssci.2018.05.015

Rashidy, R. A. H. E., Hughes, P., Figueres-Esteban, M., Harrison, C., & Van Gulijk, C. (2018). A big data modeling approach with graph databases for SPAD risk. *Safety science, 110*, 75-79. doi:10.1016/j.ssci.2017.11.019

Raviv, G., Shapira, A., & Fishbain, B. (2017). AHP-based analysis of the risk potential of safety incidents: Case study of cranes in the construction industry. *Safety science, 91*, 298-309. doi:10.1016/j.ssci.2016.08.027

Reinach, S., & Viale, A. (2006). Application of a human error framework to conduct train accident/incident investigations. *Accident Analysis & Prevention, 38*(2), 396-406. doi:10.1016/j.aap.2005.10.013

Saaty, R. W. (1987). The analytic hierarchy process—what it is and how it is used. *Mathematical Modelling, 9*(3), 161-176. doi:10.1016/0270-0255(87)90473-8

Saaty, T. L. (1988). *What is the Analytic Hierarchy Process?*, Berlin, Heidelberg.



doi:10.6342/NTU202001876

Saaty, T. L. (1990). How to make a decision: The analytic hierarchy process. *European Journal of Operational Research, 48*(1), 9-26. doi:10.1016/0377-2217(90)90057-I

Saaty, T. L. (2000). *Fundamentals of decision making and priority theory with the analytic hierarchy process* (Vol. 6): RWS publications.

Saaty, T. L. (2008). Decision making with the analytic hierarchy process. *International Journal of Services Sciences, 1*(1), 83-98. doi:10.1504/IJSSCI.2008.017590

Sen, P. K. (1968). Estimates of the Regression Coefficient Based on Kendall's Tau. *Journal of the American Statistical Association, 63*(324), 1379-1389. doi:10.1080/01621459.1968.10480934

Tabai, B. H., Bagheri, M., Sadeghi-Firoozabadi, V., & Sze, N. (2018). Evaluating the impact of train drivers' cognitive and demographic characteristics on railway accidents. *Safety science, 110*, 162-167. doi:10.1016/j.ssci.2018.03.027

Taiwan Railways Administration (TRA). (2019). *2018 Taiwan Railways Annual Report*. Taipei City: Taiwan Railways Administration.

Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis. *Nederlandse Akademie Wetenchappen Series A., 53*, 386-392. doi:10.1007/978-94-011-2546-8_20

Vaidya, O. S., & Kumar, S. (2006). Analytic hierarchy process: An overview of applications. *European Journal of Operational Research, 169*(1), 1-29. doi:10.1016/j.ejor.2004.04.028



doi:10.6342/NTU202001876

Vargas, L. G. (1990). An overview of the analytic hierarchy process and its applications. *European Journal of Operational Research, 48*(1), 2-8. doi:10.1016/0377-2217(90)90056-H

Wilcox, R. R. (2010). Robust Regression. In *Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy* (pp. 193-215). New York, NY: Springer New York.

Wilson, J. R., & Norris, B. J. (2005). Rail human factors: Past, present and future. *Applied ergonomics, 36*(6), 649-660. doi:10.1016/j.apergo.2005.07.001

Zahedi, F. (1986). The Analytic Hierarchy Process—A Survey of the Method and its Applications. *INFORMS Journal on Applied Analytics, 16*(4), 96-108. doi:10.1287/inte.16.4.96

Zhao, Y., Stow, J., & Harrison, C. (2018). A method for classifying red signal approaches using train operational data. *Safety science, 110*, 67-74. doi:10.1016/j.ssci.2017.12.007

交通部臺灣鐵路管理局(2016)。**交通部臺灣鐵路管理局行車實施要點第 280 條**。[Taiwan Railways Administration (TRA). (2016). Taiwan Railways Administration Directions for Train Operations, §280.]

交通部臺灣鐵路管理局(2018)。**交通部臺灣鐵路管理局列車自動防護系統(ATP)使用及管理要點**。[Taiwan Railways Administration (TRA). (2018). Taiwan Railways Administration Directions for Usage and Management of the Automatic Train Protection System.]



doi:10.6342/NTU202001876

行政院（2019）。**臺鐵總體檢報告**。臺北市：行政院。取自 https://event.motc.gov.tw/home.jsp?id=1986&parentpath=0,1985&websiteid=201812270001 [Executive Yuan. (2019). *Taiwan Railways Administration General Inspection Report*. Taipei, Taiwan: Executive Yuan. Retrieved from https://event.motc.gov.tw/home.jsp?id=1986&parentpath=0,1985&websiteid=201812270001]

行政院 1021 鐵路事故行政調查小組（2018）。**臺鐵 6432 次列車新馬站內正線出軌事故調查事實、原因及問題改善建議報告**。臺北市：行政院。取自 https://event.motc.gov.tw/home.jsp?id=1986&parentpath=0,1985&websiteid=201812270001 [Executive Yuan 1021 Railway Accident Administrative Investigation Team (Executive Yuan 1021 RAAIT). (2018). *Taiwan Railways Administration Train No. 6432 Derailment in the Main Line of Xinma Station Survey facts, Causes and Problem Improvement Suggestions Report*. Taipei, Taiwan: Executive Yuan. Retrieved from https://event.motc.gov.tw/home.jsp?id=1986&parentpath=0,1985&websiteid=201812270001]

孫碩昱（2010）。**鐵路司機員駕駛行為分析之研究**（碩士論文）。國立成功大學，臺南市。取自 https://hdl.handle.net/11296/tv58nt [Sun, S.-Y. (2010). *Train Drivers' Driving Behavior Investigation.* (master thesis, National Cheng Kung University, Tainan City). Retrieved from https://hdl.handle.net/11296/tv58nt]

國立臺灣大學軌道科技研究中心（2012）。**捷運萬大線風險評估模式**。 [National Taiwan University Railway Technology Research Center (NTURTRC). (2012). *Risk Assessment Models of MRT Wanda Line*.]



doi:10.6342/NTU202001876

臺灣鐵路管理局電務處（2005）。**號誌設備概要**。臺北市：臺灣鐵路管理局電務處。取自 https://www.railway.gov.tw/tra-tip-web/adr/about-publication [Taiwan Railways Administration Electrical Engineering Department (TRA Electrical Engineering Department). (2005). *Introduction of the Signal System*. Taipei, Taiwan: Taiwan Railways Administration Electrical Engineering Department. Retrieved from https://www.railway.gov.tw/tra-tip-web/adr/about-publication]

龐巴迪運輸公司（2005）。**台鐵 ATP 車上設備**。[Bombardier Transportation. (2005). *ATP On Board Equipments of TRA*.]



doi:10.6342/NTU202001876