Accident Analysis and Prevention 117 (2018) 10–20

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Exploring rainfall impacts on the crash risk on Texas roadways: A crash-based matched-pairs analysis approach

Ehsan Omranian*, Hatim Sharif, Samer Dessouky, Jose Weissmann

Department of Civil and Environmental Engineering, University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249, United States

ARTICLE INFO

Keywords:
Weather effect
Accident risk
Crash-based matched pair analysis
Road safety
Injury severity
Radar rainfall

ABSTRACT

Interaction between adverse weather conditions and motor vehicle crashes is an important topic for traffic engineers and hydrometeorologists. With the recent availability of high resolution precipitation products (hourly, 4 × 4 km), it is possible to evaluate crash risk during rainfall events more accurately. Texas, the second largest state in the U.S., with a relatively high population density in its eastern part that receives significant rainfall from tropical events, experiences many hazardous traffic conditions every year. This study investigates temporal and spatial variability of the Relative Accident Risk (RAR) due to rainy conditions across Texas during the year 2015 using a Crash-Based Matched Pairs Analysis (CB-MPA) approach for every 4 × 4 km grid using an hourly time scale. The overall findings show that rainfall increases crash risk across the state by about 57%, while seasonal-based analysis confirms the role of precipitation patterns on crash rates. Although eastern and central counties (wetter and more urbanized) have remarkably higher rates of crash occurrence, the western counties (mainly rural and dry) show higher RAR values. Moreover, higher rainfall intensity can increase RAR up to three-fold while directly having an adverse effect on crash injury type. There is a relatively high correlation between rainfall intensity and RAR values ($R^2 = 0.76$). The analysis also shows higher RAR values on high-speed interstate highways and tollways compared to urban local streets. RAR values also vary according to the gender and age of drivers. The study findings shed light on future paths toward more detailed applications of high-resolution environmental data in crash risk analysis.

1. Introduction

Roadway-related crashes are among the five leading causes of death worldwide (WMO, 2015). Adverse weather conditions are well documented as an important environmental factor leading to higher risk of motor vehicle crash occurrences (Mannering and Bhat, 2014; Tamerius et al., 2016; Peng et al., 2017). A number of studies have investigated the safety effects of different weather variables including temperature (Yu et al., 2013; Basagaña et al., 2018), fog and smog (Abdel-Aty et al., 2011), wind (Brijs et al., 2008) and precipitation (Jung et al., 2010; Bergel-Hayat et al., 2013; Black et al., 2016; Tamerius et al., 2016). Almost, all studies confirmed the significant impact of precipitation on crash risk and frequency. However, this impact has various compounding factors such as visibility impairment and road friction which were considered across several studies (Black et al., 2016; Tamerius et al., 2016; Kouchaki et al., 2017). Qui and Nixon (Qiu and Nixon, 2008) meta-analysis of previous studies indicates that solid precipitation (snow) and rainfall can increase the average relative risk of crash and injury rates by 84 and 75% respectively for snow and 71 and 49% respectively for rain. However, the wide confidence intervals of estimates based on multiple investigated studies revealed that these ranges can vary considerably. Black et al. (2016) illustrate that there is a statistically significant increase in crash (10%) and injury (8%) rates during rainy days, while higher precipitation intensities may increase the probability of crashes by up to 50%. Confirming this issue, a study by Hambly et al. (2013) demonstrated that days with moderate and heavy rainfall (≥10 mm) are associated with more crash risk.

Despite the fact that precipitation and temperature records provide a complementary source to build more accurate predictive models for crash risk analysis and distinguish their effective factors, many uncertainties make such task very challenging. Ye and Lord (2014) suggested that sample sizes may have an effect on crash risk models' output and accuracy and concluded that small sample sizes can significantly affect the development of reliable analytical models. In addition, Yu et al. (2013) and Yu and Abdel-Aty (2014) attempted to investigate the effect of adverse weather conditions using crash occurrence models. Authors suggested that limited records of fatalities to investigate the severity of crashes based on real-time weather data was one of the main

---

* Corresponding author.
E-mail addresses: seyedehsan.omranian@utsa.edu (E. Omranian), hatim.sharif@utsa.edu (H. Sharif), samer.dessouky@utsa.edu (S. Dessouky), jweissmann@utsa.edu (J. Weissmann).

https://doi.org/10.1016/j.aap.2018.03.030
Received 29 October 2017; Received in revised form 22 March 2018; Accepted 28 March 2018

concerns of the study.

Mismatch between the temporal resolutions of crashes and environmental data may cause significant underestimation or overestimation of Relative Accident (crash) Risk (RAR) (Jaroszweski and McNamara, 2014; Black et al., 2016). Both temporal and spatial resolutions of precipitation records can have significant effects on the calculation of the relative crash risk, highlighting the importance of these parameters. Various temporal scales have been used to analyze precipitation effects on RAR. Bergel-Hayat et al. (2013) used monthly/daily precipitation records, while in other studies daily precipitation records were used to investigate the effects of adverse weather conditions on crashes (Black et al., 2016; Jackson and Sharif, 2016). Hourly radar images were utilized to analyze the effects of precipitation based on sub-daily records compared to daily and monthly records (Jaroszweski and McNamara, 2014). Despite the lower level of accuracy, daily scales have been used more frequently instead of other more refined temporal scales. The main reason is the availability of daily data. However, a rainy day may have several hours with no rain or may only have a short rainy period. Neglecting these dry intervals during the daily period may cause underestimation of the impact of rainfall on RAR (Hambly et al., 2013; Jaroszweski and McNamara, 2014). On the other hand, utilizing sub-daily scales of precipitation records may falsely highlight short-term adaptations in the driver-behavior factor which inversely causes a greater impact on the risk estimates. Consequently, a more detailed approach is needed to overcome these various temporal scale deficiencies.

The spatial distribution of weather stations in the region of interest creates significant uncertainties in identifying weather conditions at the crash time and location (Tamerius et al., 2016). Since weather conditions vary spatially (particularly for large-scale study areas), sparsely distributed weather stations are not expected to provide accurate data about weather conditions at the location of each crash. This problem is more critical for rural regions usually covered by a lower density of weather stations which are typically located in highly-populated urban regions (Black et al., 2016). Therefore, for large-scale analysis of crash risks, it is not possible to identify the meteorological conditions associated with crashes in these rural areas (Bergel-Hayat et al., 2013; Jaroszweski and McNamara, 2014; Black et al., 2016; Tamerius et al., 2016). For this reason, Jaroszweski and McNamara (2014) employed radar data to provide more accurate precipitation records at crash locations. Using gridded radar data can help estimating crash risk due to rainfall at locations with no weather stations and therefore makes it possible to compare the RAR between urban and rural regions. In this study, gridded hourly precipitation data is used to investigate the large-scale effect of precipitation on crash risk in Texas. As one of the most accident-prone states in the U.S., with different climatic zones due to its large surface area, Texas is an ideal region to investigate the effects of rainfall on crashes using radar data. Moreover, seasonal climatic characteristics, close proximity to the Gulf of Mexico and an increasing population justify the need for a large-scale analysis of the interaction of crashes and rainfall over Texas. In addition, several extereme rainfall events occurred during 2015 (especially in the summer months) resulting in considerable damages (Wolter et al., 2016; Omranian and Sharif, 2018) to the roadways, bridges and other infrastructure along with a high potential for an increase on fatal crashes. With the recent advances and tools in generating flood inundation maps and flood forecasting (Afshari et al., 2016, 2018; Javaheri et al., 2018), it is possible to distinguish high-risk flood zones, however, a more complicated data-driven approach is needed to connect weather conditions to traffic data.

In this paper, we provide a comprehensive answer to this question: what would be the large-scale effect of adverse weather conditions on crash occurance?. In another part of the paper, a background about the applied analysis method and its inner workings is broadly explained. A detailed information about the study area, datasets (crash/precipitation) and the modefied analysis methodology are also provided. Finally, results and discussion are expanded together with conclusions.

2. Background: matched-pair analysis method

2.1. Matched-pair analysis application

One of the most challenging issues related to crash analysis and development of crash risk statistical models is the identification of the time-dependent variables that have a strong influence on crash rates (Andrey, 2010; Sun et al., 2011; Omranian et al., 2018). Researchers have identified many confounding factors that can increase the uncertainty of crash risk analysis (Datla and Sharma, 2008; Black et al., 2016). To decrease the uncertainty due to confounding factors, the Matched-Pairs Analysis (MPA) was used to investigate the impact of adverse weather conditions on crash rates and severity (Andrey and Yagar, 1993; Andrey et al., 2003; Eisenberg, 2004; Keay and Simmonds, 2006; Andrey, 2010; Mills et al., 2011; Sun et al., 2011; Jaroszweski and McNamara, 2014; Black et al., 2016; Tamerius et al., 2016). The remarkable characteristics of confounding factors such as traffic volume are their fluctuations across time (hour of a day, day of the week, month or season) and space. Therefore, the idea behind MPA is to use corresponding control periods for the same location with reported crashes to eliminate the impact of time-sensitive factors such as traffic volume, traffic pattern, driver behavior, light conditions, landuse density and road construction. For instance, traffic volume and patterns are assumed to be generally similar during same hours of the same weekdays over time (excluding specific holidays).

2.2. Time-Step selection criteria

MPA may select matched pairs based on different time intervals which are generally separated by one week. Due to data limitations and complexity of matching algorithms needed for hourly analysis, only a few studies used hourly matched-pair analysis (Tamerius et al., 2016). However, three-hourly (Keay and Simmonds, 2005; Jaroszweski and McNamara, 2014), six-hourly (Andrey et al., 2003; Andrey, 2010; Mills et al., 2011) and daily (Eisenberg, 2004; Black et al., 2016) analysis were applied in other studies. Moreover, other researchers (Sun et al., 2011) used flexible intervals of up to 12 h based on rainfall duration and considered control dry periods with the same length.

In a daily analysis approach, many crashes may be falsely considered as wet crashes because they occurred in a day which has observed precipitation above a specific threshold even if the rainfall lasted for a short time. Therefore, all crashes that occurred during dry hours would be counted as wet crashes. The higher temporal resolution gains more importance in regions such as Texas with frequent severe thunderstorms characterized by very high rainfall intensities of short duration and limited spatial coverage (Schlögl and Stütz, 2017; Furl et al., 2018; Omranian and Sharif, 2018).

For analysis under an hourly time-step, there is no difference between rainfall intensity and cumulative precipitation amount. Therefore, each hour with any amount of precipitation record may be considered as a “wet-hour”. However, in greater time-scales (e.g., six-hourly or daily), there should be a minimum threshold of cumulative precipitation records during a time period to be considered as a “wet-period”. For example, Andrey (2010) defined a six-hour wet-period when the equivalent liquid precipitation amount was greater than 0.4 mm while Black et al. (2016) used a 0.254 mm cumulative precipitation amount as a limit for recognizing a day as a “wet-day”.

2.3. Relative accident risk (RAR) computation

In an hourly analysis approach, if wet and dry period occurs on a specific hour of a day with exactly one-week interval at the same geographic location, respectively, then a matched pair is identified. Matched pair selection is first based on next following week. If a