PROACTIVE SAFETY ASSESSMENT OF URBAN THROUGH-ROADS BASED ON GPS DATA

Jiří AMBROS1, Jan ELGNER2, Veronika VALENTOVÁ3, Radosław BĄK4, Mariusz KIEĆ5

1, 2, 3 CDV – Transport Research Centre, Brno, Czechia
4, 5 Faculty of Civil Engineering, Cracow University of Technology, Cracow, Poland

Abstract:
Road safety is a worldwide issue, while urban roads account for a high share of serious road injuries, especially involving vulnerable road users, such as pedestrians or cyclists. Specifically, the safety of major roads through built-up areas (through-roads) is insufficient due to mixed traffic conditions including vulnerable road users, varying driving behaviour, and many disruptions, which are combined with excessive speed. In this context, various traffic calming measures have been implemented to improve road safety, such as gateways or pedestrian refuge islands. However, the specific safety impacts of traffic calming combined with specific characteristics of through-roads are often unknown, since most traditional evaluations have been limited by small sample sizes of crash data, as well as wide variations in physical and road characteristics. To overcome the limitations of crash-based evaluations, we used the GPS-based data from a sample of 21 Czech and 12 Polish through-roads to develop the Speed-Safety Index, which combines speed, speed variance, and traffic volume. Our study has three novelty features: (1) To assess safety, we used speed and speed variance simultaneously. (2) To complete the missing link between specific traffic calming measures and safety, we validated the statistical relationship between the developed Speed-Safety Index and crash history. (3) To prove the usefulness of the developed index, we also showed its practical interpretation by proving the effect of spacing between traffic calming measures on safety. The index proved to be well correlated to crash frequency and it also proved the effect of spacing between traffic calming measures: the longer spacing, the smaller speed-reducing effect.

The paper concludes with a discussion on the limitations, which we plan to address in further research, by moving from the current macro-perspective (Speed-Safety Index on the level of through-roads) to the micro-perspective (focusing on individual directions, locations, and traffic calming measures). We also plan to investigate the method’s applicability in different contexts. If the approach proves feasible, with reliable and valid results, it may become an alternative for a proactive network-wide road assessment, as called for by the European Road Infrastructure Safety Management Directive.

Keywords: road safety, urban road, traffic calming, speed

To cite this article:

Contact:
1) jiri.ambros@cdv.cz [https://orcid.org/0000-0003-2707-6243] – corresponding author; 2) jan.elgner@cdv.cz [https://orcid.org/0000-0002-6086-6565]; 3) veronika.valentova@cdv.cz [https://orcid.org/0000-0003-0938-4429]; 4) rbak@pk.edu.pl [https://orcid.org/0000-0002-3115-3486]; 5) mkiec@pk.edu.pl [https://orcid.org/0000-0001-9193-7269]

Article is available in open access and licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0)
1. Introduction

Road safety is a worldwide issue; although the most severe crashes occur on rural roads, urban roads and streets account for a high share of serious road injuries (OECD/ITF, 2022). Urban roads are particularly critical for vulnerable road users (VRUs), such as pedestrians or cyclists, who represent more than 70% of road fatalities in cities (Adminaitė-Fodor and Jost, 2019).

In Central and Eastern Europe, major roads through built-up areas (through-roads) pose a specific problem (Gaca and Pogodzińska, 2017; Ambros et al., 2021a). Although they carry a dominant share of through traffic (in addition to local traffic) and significant traffic volumes, they often have no hierarchical structure and lack access control, which can cause traffic disruptions and as a result, speed changes. These conditions, together with the mentioned presence of VRUs, are then reflected in insufficient safety performance. In this context, the current European directive on road infrastructure safety management (RISM; EU, 2019) recommends improving the safety of VRUs by considering their needs in all RISM procedures and developing quality requirements for VRU infrastructure, as well as introducing a proactive approach in safety evaluation.

In fact, all these issues are closely related to driving speed. Speed has been recognized as the most influential risk factor (OECD/ITF, 2018), contributing to around 30% of all fatal crashes (Ambros et al., 2020; EC, 2020; Van den Berghe, 2021; NHTSA, 2023; Soole et al., 2023). Specifically, on Czech and Polish roads, speeding has been attributed to approximately 40% of fatal crashes in recent years, making it the most frequent cause of road deaths (Straka and Pelešková, 2023; Symon and Rzepka, 2023). In this context, various physical speed management measures (traffic calming measures, TCMs) have been implemented, such as road narrowing, pedestrian refuge islands, or speed cameras to provide more fluent traffic with lower speed and dispersion. However, the specific safety impacts of traffic calming combined with specific characteristics of through-roads (mixed traffic and a number of disruptions) are often unknown, since most evaluations have been limited by small sample sizes of crash data, as well as wide variations in TCM and road characteristics (more details are listed in Section 2).

To sum up, through-roads are not sufficiently safe due to inappropriate design and speeds, in combination with the mixed traffic, the increased presence of VRUs and through traffic, including heavy vehicles. Although various TCMs are available, their safety impacts are often not well known, which complicates the systematic application of TCMs combined with specific characteristics of through-roads. Crashes are random events, which are statistically rare, as well as often underreported; they are also not very informative about the exact causes of crashes, and their use raises ethical concerns, since one has to wait for crashes to occur, and thus for people to suffer, before the road safety situation can be evaluated (Lord et al., 2021). To overcome the limitations of traditional crash-based evaluations, we decided to assess the safety based on the speed data. Specifically, GPS data were used, which allows obtaining speed and its variation along the entire analysed through-roads, without being limited to any stationary measurements. As a result, we developed so called Speed-Safety Index (SSI) for proactive safety assessment. More background information is provided in the following section, followed by Data and methods, Results, Discussion and Conclusions.

2. Background and objectives

The following literature review provides more background information on the main elements of the presented study: traffic calming, and speed and speed variance. The section concludes with a summary and description of the paper’s novelty.

2.1. Traffic calming

To reduce driving speed and increase safety, traffic calming has been applied in Western Europe since the 1960s. Generally, it provided measurable safety benefits, followed by positive feedback and a worldwide uptake, including in Australia (1980s), Northern America (1990s), and Central/Eastern Europe (2000s).

Surprisingly, despite the long history of traffic calming, the safety impacts of TCMs are often uncertain or unknown (Ambros et al., 2023). Several studies focused on the impact of TCMs on crashes. For example, a Cochrane systematic review (Bunn et al., 2003) noted that “further rigorous evaluation is needed” to determine the effectiveness of TCMs. An Australian study (Sobhani et al., 2016), which attempted to summarize the effects of TCMs, admitted...
limited sample sizes and less reliable results. An international synthesis by Yannis et al. (2015) found injury crash reductions from traffic calming varying widely between 8 and 50%.

The possible reason for the uncertain evaluations and varying findings may be that TCMs are widely varying in their design, configuration, or surrounding conditions, which complicates assessing their general effectiveness. Additionally, the spacing, i.e., the distance between TCMs, plays a role (Moreno and García, 2013). This is linked to the fact one of the key goals of traffic calming is to create a consistently lower speed environment along a route or across a whole urban area, rather than to cause localized speed reduction (Brindle, 2005). To achieve this, instead of isolated elements, a series of TCMs with specific spacings are required to moderate speed.

2.2. Speed and speed variance

Given the clear relationship between speed and safety (increasing frequency and severity of crashes (OECD/ITF, 2018), speed is a potential safety indicator. Nevertheless, it was indicated that also speed variability plays a role (Aarts and van Schagen, 2006; OECD/ITF, 2018). Note that there are two approaches to defining the speed dispersion – either as speed variation between individual vehicles, or speed variance at the road section level – while in the further text we focus on the latter.

Speed variance is related to road and traffic (and in turn to driving behaviour) in several perspectives:

- Increased driving difficulty and greater speed variance were found in locations, where changes to the visual appearance of the road were made, without any changes to the road geometry (Charlton and Starkey, 2013). Driving workload and road complexity were found to increase speed variance (Edquist et al., 2012); the work zone is an illustrative example of such an environment (Steinbakk et al., 2019).

- Speed variance may indicate less predictable road design, which leads to poor driver expectancy and lack of consistency in driver behaviour (Farah et al., 2017).

- Dense traffic often leads to frequent and sudden changes in speed, related to lane-changing and overtaking (Dell’Acqua, 2011; ) and thus to increased speed variance.

- Speed variance may also reflect how drivers perceive the road and speed limits; in other words, the extent to which the road is self-explaining. When comparing speed perception in various road sections, Ambros et al. (2021b) found that on rural roads the participants chose similar driving speeds, probably due to their relative uniformity; in contrast to urban and transition sections with more mixed characteristics, where speed variability was larger.

Interestingly, findings on the relationship between speed variance and crash risk have been very diverse. While Aarts and van Schagen (2006) claimed that larger speed variance is related to a higher crash rate, Pei et al. (2012) found speed variation not significantly associated with crashes. The differences may be due to data aggregations:

- For example, Gitelman et al. (2017) found the speed variance related negatively to daytime crashes and positively to nighttime crashes.

- When developing prediction models, Figueroa-Medina and Tarko (2005) identified several speed dispersion factors related to road geometry and cross-section; however, they were different between tangents and curves.

- Shinar (2017) found speed deviations positively related to crashes on rural roads, but not on urban roads.

All the mentioned aggregations may bias the findings and thus mask the true associations between speed variance and crash risk.

As described, speed variance reflects several dimensions (e.g., workload, complexity, predictability) which may be linked to through-roads and TCMs. However, most of the mentioned studies were conducted in rural settings. This is probably why several traffic calming manuals/reviews (Brindle, 2005; Juriewicz, 2009; Hillier et al., 2016) reported the impacts on speed, but not on speed variance. In addition, some studies, which investigated the impacts of TCMs on speed variance (Mountain et al., 2005; Daniel et al., 2011; Agerholm et al., 2017) found relatively small effects which did not allow clear conclusions.

In summary, while speed variance theoretically provides interesting information, in practice the findings are very diverse and its relationship to crashes is mixed. Even the relationship to speed is not fully clear; while some found a positive tendency (higher speed associated with a larger variance; Figueroa-
Medina and Tarko, 2005), others found the opposite (higher speed associated with a lower variance; Aarts and van Schagen, 2006). As summarized by Dell’Acqua (2011), “certain engineering measures might reduce mean traffic speeds but at the same time increase the speed variability to an extent at which the accident frequency may stay the same or even rise.”

To obtain speed-related data, GPS has been used since the 2010s. Nevertheless, we are aware of just one study, which attempted to relate GPS-based speed indicators to crashes: Moreno and García (2013) developed two indices as surrogate measures of safety based on GPS data and used them for comparison of scenarios with different spacing between TCMs. However, they could not investigate the relationship between the indexes and crashes due to low crash frequencies.

2.3. Summary and novelty

Based on the literature review, we decided to contribute to the state-of-the-art as follows:

1. To assess safety, we used speed and speed variance simultaneously.
2. To complete the missing link between specific TCMs and safety, we validated the statistical relationship between the developed Speed-Safety Index and crash history.
3. To prove the usefulness of the developed index, we also showed its practical interpretation by proving the effect of spacing between traffic calming measures on safety.

3. Data and methods

Through-roads are not sufficiently safe due to inappropriate design and speeds, in combination with the increased presence of VRUs, through-traffic including heavy vehicles and possible disruptions. Therefore road safety assessment of through-roads is difficult. At the same time, road traffic safety may be improved by traffic calming. Although various traffic calming measures are available, their safety impacts are often not well known, which complicates the systematic application of measures. However, this limitation may stem from the fact that previous studies usually used traditional crash-based evaluations. In fact, to evaluate the effects of the road- and speed-related treatments, it would be more straightforward to use the road or speed data; in other words: focusing on inputs and outputs (the behaviour and its determinants) instead of the final outcome (safety in terms of crashes). Fig. 1 visualizes the hypothesized relationship between road and roadside – or road(side) – characteristics, behaviour and safety, together with their example operationalizations. This motivated us to develop the safety assessment of urban through-roads based on quantitative analysis using GPS-based behavioural data (speed and speed variance) to develop the Speed-Safety Index (SSI).

3.1. GPS data

The approach to collecting GPS data was different for the Czech and Polish samples. In Czechia, GPS was retrieved from the already existing database from a fleet of commercial probe vehicles, collected at a 0.25-second rate (4 Hz frequency) for 8 months (Oct 2014 – May 2015); for more details see Ambros et al. (2017). The selection focused on through-roads including physical elements (mainly gateways and pedestrian refuge islands) and excluding any other potentially disrupting elements (e.g., signalized intersections, roundabouts, railway level crossings) and with frequent drives presented in the database. The area of interest was 200 m before and after the urban area limits (traffic signs). In total 21 through-roads were selected, with lengths between 0.3 and 2.8 km, involving 15 to 309 drives in each direction. Traffic volume based on the 2016 national traffic census was between approx. 3,000 and 15,000 veh/day. The speed limit was 50 km/h.

![Fig. 1. The hypothesized relationship and example operationalizations](image)
In Poland, the floating car technique was used: following randomly selected leading vehicles traveling at the free-flow speed at a constant gap with an instrumented vehicle, collecting data at a 0.1-second rate (10 Hz frequency) together with a video record; for more details see Gaca and Kieć (2016). Data collection was conducted in 2015–2016 during the day and in good weather conditions. Each drive started and finished outside a built-up area. In total 12 through-roads with physical elements such as gateways and pedestrian refuge islands, with lengths from 0.7 km to 7.2 km, were driven this way in both directions (10 to 25 drives in each direction). The selected sites were relatively homogeneous in terms of cross-section, road surroundings, lack of access control, and only a small share of local traffic. Traffic volume based on the 2015 national traffic census (annual average daily traffic, AADT) varied between 5,000 and 18,500 veh/day. The speed limit was 50 km/h.

**Czech examples:**

![Jaroslav (road I/35)](image)

![Lomnice nad Lužnicí (road I/24)](image)

**Polish examples:**

![Bibice (road DK7)](image)

![Niedzwica Duża (road DK19)](image)

Fig. 2. Example photographs of through-roads from Czech and Polish samples

Example photographs from the selected roads in both Czech and Polish samples are in Fig. 2. Basic data (AADT and length) is provided in the summary table in the Appendix.

For illustration, Fig. 3 presents an example of a speed-distance graph (also known as speed profile) for both driving directions, together with their confidence intervals. The dashed lines indicate the location of potential speed influencers (village limits and traffic calming measures).

Note that in the following analyses we used speed and speed variance from the entire through-roads and both driving directions. The summary table in the Appendix provides speed and speed variance for each through-road.

### 3.2. Crash data

The frequency of Police reported crashes was assigned to each through-road. Note that they were not split into driving directions because default crash data does not include such detailed information.
In the Czech sample, 10 years (2010–2019) of both property-damage-only (PDO) and injury crash data was retrieved from the Traffic Police database. In Poland, where PDO crashes are not routinely registered, only injury crash data could be used (11-year period 2006–2016). In the following text, “all crashes” (sum of PDO and injury crashes) are reported for the Czech sample, while “injury crashes” are reported for the Polish sample.

The summary table in the Appendix provides the annual average number of injury crashes.

### 3.3. Speed-Safety Index

The starting point of developing a speed-related safety index was the relationship between three dimensions of road safety, as described by OECD (1997): the first is the magnitude of the activity that results in accidents (exposure), the second dimension is the accident risk situation, and the third dimension is the accident consequence (severity of accident). These three dimensions cover the three main effects on safety: change in any one of these dimensions changes the entire safety level.

\[
\text{Safety} = \text{Exposure} \times \text{Risk} \times \text{Consequence} \tag{1}
\]

We proposed a Speed-Safety Index (SSI) for a chosen road segment, with three elements: speed, standard deviation of speed, traffic volume. These elements are related to formula (1) as follows:

- **Risk** is represented by \( \sigma_i \) – standard deviation of speed at the \( i \)-th segment (larger variability of speed should be associated with higher risk).
- **Consequence** is represented by \( S_i \) – the speed at the \( i \)-th segment (higher speed should be associated with higher severity of accidents).
- **Exposure** is represented by the traffic volume at the \( i \)-th segment (AADT).

For a segment of length \( L \), the SSI is defined as follows:

\[
SSI = \int_0^L \sigma \cdot S \cdot AADT \cdot dL \tag{2}
\]

In practice, SSI may be calculated as the sum of the products of speed and standard deviation with the length interval, depending on the frequency and accuracy of GPS data (frequency of collecting GPS speed data), multiplied by AADT, which is constant along the segment:

\[
SSI = AADT \cdot \left( \sum_{i=1}^{n} \sigma_i \cdot S_i \right) \tag{3}
\]

where \( n \) is the number of segments. Formula 3 was used in the following analysis.

The formula for calculating the SSI is a function that relates the variables of traffic volume (AADT), road...
section length (number of constant length segments), and other factors affecting road safety, characterized by driver behaviour (speed and speed dispersion).

\[
SSI = f(AADT, \text{length, road characteristics}) \tag{4}
\]

This approach is commonly used in road safety analyses, e.g., the Highway Safety Manual (AASHTO, 2010), where the safety performance function also considers the impact of AADT, section length, and road infrastructure characteristics (road design, traffic calming measures, roadside, driveways, etc.) on the number of crashes. In the case of road safety assessment by SSI, the impact of road infrastructure characteristics is characterized by the behaviour of drivers (speed and its dispersion).

Table 1 presents descriptive statistics of all variables in the through-road sample used in the analyses of the Speed-Safety Index.

### 4. Results

The speed and speed variance data were used to calculate SSI according to formula 3, assuming a constant segment length of 10 m. For better display, a fraction of SSI divided by \(10^6\) is further used. Final data for each through-road is provided in the summary table in the Appendix. The relationship between SSI and annual average crash frequency was assessed by Pearson correlation. Similarly, as in Section 4.2, the frequency of “all crashes” was available for the Czech sample; the Polish sample consisted only of injury crashes. The results are listed in Table 2. All correlation coefficients were statistically significant with \(\geq 95\%\) confidence (\(\leq 0.05\) significance level) and exceeded 0.8, which indicates a high correlation (Hinkle et al., 2003). The correlation is positive, which means that increasing SSI is associated with more crashes.

For interpretation, we wanted to use some aggregated indicator of traffic calming performance – we used the average spacing between the TCMs, such as gateways and pedestrian refuge islands. We calculated the Pearson correlation of spacing to the SSI, as well as crash frequency. As previously noted, the Czech data allowed using injury crashes and all crashes, while the Polish sample included injury crashes only. The results are provided in Table 3.

### Table 1. Descriptive statistics of data used in the analysis

<table>
<thead>
<tr>
<th></th>
<th>AADT [veh/day]</th>
<th>Length [km]</th>
<th>Average speed [km/h]</th>
<th>Coefficient of variation</th>
<th>Injury crashes / year</th>
<th>Injury crash rate / 10^6 veh-km</th>
<th>SSI / 10^6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Czech sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum</td>
<td>731</td>
<td>0.3</td>
<td>44.91</td>
<td>0.09</td>
<td>0.0</td>
<td>0.0</td>
<td>48.0</td>
</tr>
<tr>
<td>Maximum</td>
<td>15,367</td>
<td>2.8</td>
<td>58.85</td>
<td>0.21</td>
<td>1.9</td>
<td>2.54</td>
<td>2,777.1</td>
</tr>
<tr>
<td>Average</td>
<td>8,197</td>
<td>1.1</td>
<td>52.24</td>
<td>0.13</td>
<td>0.7</td>
<td>0.34</td>
<td>509.6</td>
</tr>
<tr>
<td><strong>Polish sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum</td>
<td>763</td>
<td>0.7</td>
<td>44.31</td>
<td>0.11</td>
<td>0.0</td>
<td>0.0</td>
<td>556.5</td>
</tr>
<tr>
<td>Maximum</td>
<td>16,505</td>
<td>7.2</td>
<td>65.15</td>
<td>0.25</td>
<td>10.0</td>
<td>0.43</td>
<td>7,922.3</td>
</tr>
<tr>
<td>Average</td>
<td>9,290</td>
<td>3.0</td>
<td>54.35</td>
<td>0.17</td>
<td>2.6</td>
<td>0.18</td>
<td>2,654.5</td>
</tr>
</tbody>
</table>

### Table 2. Correlations between SSI and crashes

<table>
<thead>
<tr>
<th>Pearson coefficient of correlation between SSI and…</th>
<th>Czech sample</th>
<th>Polish sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>… injury crashes</td>
<td>0.82</td>
<td>0.83</td>
</tr>
<tr>
<td>… all crashes</td>
<td>0.92</td>
<td>–</td>
</tr>
</tbody>
</table>

### Table 3. Correlations between spacing, SSI and crashes

<table>
<thead>
<tr>
<th>Pearson coefficient of correlation between spacing and…</th>
<th>Czech sample</th>
<th>Polish sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>… SSI</td>
<td>0.80</td>
<td>0.56</td>
</tr>
<tr>
<td>… injury crashes</td>
<td>0.66</td>
<td>0.62</td>
</tr>
<tr>
<td>… all crashes</td>
<td>0.75</td>
<td>–</td>
</tr>
</tbody>
</table>
The correlation coefficients in the Czech sample were around or above 0.7, which indicates a high correlation; in the Polish sample, they were around 0.6, which indicates a moderate correlation (Hinkle et al., 2003). All correlation coefficients were statistically significant with $\geq 90\%$ confidence ($\leq 0.10$ significance level). The correlation is positive, which confirms that increasing spacing (i.e., a longer distance between TCMs) is associated with a smaller traffic calming effect and thus lower safety (higher SSI and more crashes).

5. Discussion
The developed SSI seems to be a very promising indicator that may be used to assess road safety on various homogeneous sections of road infrastructure (rural roads, through-roads, etc.) based on GPS data. This is a proactive approach, which enables the identification of hazardous sections in the road network without analyzing the crash databases. Nevertheless, the SSI also has some limitations:

- SSI includes section length, which means that longer sections may result in a higher number of crashes. To reduce this impact, the SSI can be divided by the overall section length to obtain the SSI per km value.

- The value of SSI depends on the interval length of the segment (10m in current analyses) from which the data is summed up. If the section is smaller, higher SSI is obtained, but at the same time, speed variability is analyzed more precisely. Therefore, it is necessary to use the same interval length for all analyzed sections.

- If we compare sections with similar AADT values and exclude the impact of length (by using the SSI per km value, mainly speed and its variability will have an impact on SSI.

- With comparable AADT and average speeds on road sections, the main factor influencing the SSI value is the speed variability, which may be caused by disruptions in traffic caused by mixed local and through traffic, different types of vehicles, presence of VRUs, road infrastructure (traffic calming measures, driveways, bus stops, etc.).

To illustrate the above-mentioned observations, a comparison of two Czech road sections with similar AADT and section lengths, but different speed profiles and the resulting SSI, is presented in Fig. 4.

The presented examples indicate that not only AADT and segment length (which are comparable in both cases) influence the final outcomes. The SSI in the second case (Hněvkov) are approximately 1.4 times higher, which may be due to higher speed and greater speed variation, possibly also due to lack of TCMs. This difference is also confirmed by the increased number of injury crashes.

6. Conclusions
The safety of urban roads is insufficient, due to varying driving behaviour resulting from the influence of the road environment (mixed traffic, urban and rural characteristics, different densities of buildings, driveways, etc.). For the safety assessment of through-roads, proactive (non-crash-based) approaches are needed. At the same time, speed and speeding are known to have major safety impacts. Using the GPS data from a sample of 21 Czech and 12 Polish through-roads, we developed the Speed-Safety Index (SSI), which combines speed, speed variance, and traffic volume. SSI values proved to be well correlated to crash frequency. In addition, they proved the effect of spacing between traffic calming measures: the longer spacing, the smaller speed-reducing effect.

The primary use of SSI should be road network screening to identify sections based on GPS data that have the highest risk, therefore SSI can be used to rank dangerous sections in the road network. Nevertheless, we are aware of some limitations of the applied approach:

- Although GPS data presents a valuable emerging big data source, it also has limitations, for example varying sampling rate, uncertain estimation of free-flow speed, or unknown generalizability to driving population (for a review, see Ambros et al., 2021c).

- In this study, two approaches to GPS data collection were used (probe vehicle fleet data in Czechia and floating car with video recording in Poland). While the presented results of both approaches indicate the possibility of obtaining promising results, each approach has its own (dis)advantages: the former provides big data in terms of covered time and space; the latter is demanding in terms of collection, which limits its coverage but can reveal details unseen in probe data, such as congestions or crossing pedestrians. Nevertheless, consistency between
the findings of both approaches should be verified.

When calculating SSI, driving directions were not considered independently. This aggregation may lead to overlooking potential differences between directions. The same holds for directional traffic volumes, aggregated in AADT. Also, crash data were aggregated, which may mask the differences between crash participants, configurations, or severity levels.

We plan to address these points in further research, by moving from the current macro-perspective (SSI on the level of through-roads) to the micro-perspective (focusing on individual directions, locations, and TCMs). This will require testing the impact of analysed segment length on results, as well as estimating influence zones of TCMs, the effect of segment characteristics (e.g., curvature), etc. Sensitivity tests should also include the impact of different GPS recording frequencies.

In the next steps, we plan to investigate the method’s applicability in different contexts, such as transition zones (between rural and suburban roads), and rural or high-speed roads. The SSI may perform differently in different conditions, including the impact of traffic congestion or different traffic calming measures. If the approach proves feasible, with reliable and valid results, it may become an alternative for a proactive network-wide road assessment, as called for by the European RISM directive (EU, 2019). This would benefit from ubiquitous big data coverage, including future data from autonomous vehicles, which will gradually increase coverage and reduce costs.

Fig. 4. Illustrative comparison of two Czech through-road sections
Appendix: Summary table of analysed data

<table>
<thead>
<tr>
<th>Through-road</th>
<th>AADT [veh/day]</th>
<th>Length [km]</th>
<th>Average speed [km/h]</th>
<th>Coefficient of variation</th>
<th>Injury crashes / year</th>
<th>Injury crash rate / 10^6 veh-km</th>
<th>SSI / 10^6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Czech sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jaroslav</td>
<td>14,193</td>
<td>0.4</td>
<td>54.58</td>
<td>0.14</td>
<td>0.7</td>
<td>0.31</td>
<td>121.0</td>
</tr>
<tr>
<td>Vysoká u Holic</td>
<td>13,567</td>
<td>0.3</td>
<td>53.06</td>
<td>0.17</td>
<td>0.7</td>
<td>0.54</td>
<td>48.0</td>
</tr>
<tr>
<td>Ostřetín</td>
<td>13,567</td>
<td>2.2</td>
<td>53.59</td>
<td>0.11</td>
<td>1.9</td>
<td>0.17</td>
<td>2,148.4</td>
</tr>
<tr>
<td>Chvojenec</td>
<td>13,118</td>
<td>1.2</td>
<td>51.96</td>
<td>0.13</td>
<td>1.6</td>
<td>0.28</td>
<td>676.9</td>
</tr>
<tr>
<td>Býšť</td>
<td>13,362</td>
<td>0.6</td>
<td>53.22</td>
<td>0.12</td>
<td>0.5</td>
<td>0.16</td>
<td>226.1</td>
</tr>
<tr>
<td>Malá Skála</td>
<td>4,975</td>
<td>1.3</td>
<td>50.88</td>
<td>0.11</td>
<td>0.6</td>
<td>0.25</td>
<td>285.7</td>
</tr>
<tr>
<td>Vojnův Městec</td>
<td>5,806</td>
<td>1.2</td>
<td>52.29</td>
<td>0.14</td>
<td>0.5</td>
<td>0.20</td>
<td>364.6</td>
</tr>
<tr>
<td>Hlado</td>
<td>4,649</td>
<td>1.5</td>
<td>58.85</td>
<td>0.12</td>
<td>0.4</td>
<td>0.16</td>
<td>491.6</td>
</tr>
<tr>
<td>Kámen</td>
<td>6,465</td>
<td>0.6</td>
<td>44.91</td>
<td>0.15</td>
<td>0.0</td>
<td>0.00</td>
<td>67.7</td>
</tr>
<tr>
<td>Obrataň</td>
<td>5,652</td>
<td>1.2</td>
<td>54.52</td>
<td>0.13</td>
<td>0.3</td>
<td>0.12</td>
<td>347.0</td>
</tr>
<tr>
<td>Starý Pelhřimov</td>
<td>8,467</td>
<td>0.5</td>
<td>51.72</td>
<td>0.12</td>
<td>0.1</td>
<td>0.06</td>
<td>84.0</td>
</tr>
<tr>
<td>Krahalci</td>
<td>2,735</td>
<td>1.1</td>
<td>49.93</td>
<td>0.13</td>
<td>0.4</td>
<td>0.36</td>
<td>112.8</td>
</tr>
<tr>
<td>Nová Ves</td>
<td>10,951</td>
<td>1.1</td>
<td>51.77</td>
<td>0.13</td>
<td>1.4</td>
<td>0.31</td>
<td>572.8</td>
</tr>
<tr>
<td>Vladislav</td>
<td>6,124</td>
<td>1.9</td>
<td>49.42</td>
<td>0.14</td>
<td>1.2</td>
<td>0.28</td>
<td>802.8</td>
</tr>
<tr>
<td>Lubeneč</td>
<td>7,310</td>
<td>2.8</td>
<td>57.63</td>
<td>0.14</td>
<td>1.9</td>
<td>2.54</td>
<td>2,777.1</td>
</tr>
<tr>
<td>Prácheň</td>
<td>6,916</td>
<td>1.2</td>
<td>51.69</td>
<td>0.14</td>
<td>0.8</td>
<td>0.27</td>
<td>381.2</td>
</tr>
<tr>
<td>Holohlavy</td>
<td>15,367</td>
<td>0.5</td>
<td>50.51</td>
<td>0.12</td>
<td>0.4</td>
<td>0.14</td>
<td>133.8</td>
</tr>
<tr>
<td>Voleč</td>
<td>4,692</td>
<td>1.2</td>
<td>54.30</td>
<td>0.12</td>
<td>0.3</td>
<td>0.15</td>
<td>265.1</td>
</tr>
<tr>
<td>Tlumačov</td>
<td>5,252</td>
<td>1.7</td>
<td>50.14</td>
<td>0.09</td>
<td>0.7</td>
<td>0.21</td>
<td>368.1</td>
</tr>
<tr>
<td>Hněvkov</td>
<td>6,341</td>
<td>0.5</td>
<td>50.57</td>
<td>0.21</td>
<td>0.4</td>
<td>0.34</td>
<td>95.6</td>
</tr>
<tr>
<td>Nový Dražejov</td>
<td>9,208</td>
<td>1.0</td>
<td>51.56</td>
<td>0.12</td>
<td>0.7</td>
<td>0.21</td>
<td>331.1</td>
</tr>
<tr>
<td>Polish sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bełżec</td>
<td>7,630</td>
<td>3.6</td>
<td>59.87</td>
<td>0.11</td>
<td>0.8</td>
<td>0.08</td>
<td>1,155.4</td>
</tr>
<tr>
<td>Dęblin</td>
<td>8,755</td>
<td>3.7</td>
<td>50.61</td>
<td>0.15</td>
<td>3.3</td>
<td>0.27</td>
<td>1,176.4</td>
</tr>
<tr>
<td>Niedrzwica Duża</td>
<td>13,960</td>
<td>1.8</td>
<td>53.54</td>
<td>0.20</td>
<td>0.8</td>
<td>0.08</td>
<td>1,425.0</td>
</tr>
<tr>
<td>Strzeszkowice</td>
<td>13,828</td>
<td>1.0</td>
<td>65.15</td>
<td>0.14</td>
<td>0.4</td>
<td>0.08</td>
<td>781.6</td>
</tr>
<tr>
<td>Wilkołaz</td>
<td>14,107</td>
<td>0.7</td>
<td>59.45</td>
<td>0.15</td>
<td>0.2</td>
<td>0.05</td>
<td>556.5</td>
</tr>
<tr>
<td>Wólka</td>
<td>16,505</td>
<td>1.4</td>
<td>58.60</td>
<td>0.13</td>
<td>0.0</td>
<td>0.00</td>
<td>1,025.6</td>
</tr>
<tr>
<td>Opolczno</td>
<td>15,513</td>
<td>7.2</td>
<td>48.69</td>
<td>0.15</td>
<td>10.0</td>
<td>0.24</td>
<td>7,922.3</td>
</tr>
<tr>
<td>Zabierzów</td>
<td>15,552</td>
<td>4.7</td>
<td>44.31</td>
<td>0.24</td>
<td>5.3</td>
<td>0.20</td>
<td>6,319.3</td>
</tr>
<tr>
<td>Ropczyce</td>
<td>18,470</td>
<td>2.3</td>
<td>44.90</td>
<td>0.25</td>
<td>6.7</td>
<td>0.43</td>
<td>4,103.0</td>
</tr>
<tr>
<td>Biecz</td>
<td>9,388</td>
<td>4.7</td>
<td>53.49</td>
<td>0.18</td>
<td>0.6</td>
<td>0.04</td>
<td>4,587.5</td>
</tr>
<tr>
<td>Olkusz</td>
<td>8,858</td>
<td>2.0</td>
<td>57.86</td>
<td>0.14</td>
<td>1.5</td>
<td>0.23</td>
<td>1,650.4</td>
</tr>
<tr>
<td>Mielec</td>
<td>4,963</td>
<td>2.4</td>
<td>55.75</td>
<td>0.15</td>
<td>1.8</td>
<td>0.41</td>
<td>1,151.1</td>
</tr>
</tbody>
</table>
Acknowledgements
The paper was produced with the financial support of the Czech Ministry of Transport within the programme of the long-term conceptual development of research institutes.

The help of Jan Kubeček, Richard Turek and Robert Zůvala (CDV – Transport Research Centre) with data preparation and analysis is acknowledged.

References


