The Swedish Vehicle Speed Index on National Roads

Gösta Forsman and Stig Danielsson

An important cause of traffic fatalities is speeding. In Sweden, large resources are therefore allocated to improving the observance of speed limits. To learn whether these efforts are succeeding, the Swedish National Road Administration has developed a speed index. The index is based on data from measurement sites originally designed for other purposes. Although this limits its usefulness to some extent, it does help to reduce the budget. This paper deals with the construction and interpretation of the index.

Key words: Vehicle speed; index; bootstrap.

1. Introduction

The Swedish National Road Administration (SNRA) administers the construction and maintenance of Swedish national roads. In addition, the SNRA has the overall responsibility for road traffic safety in Sweden.

Fatalities in Swedish Road Traffic fell from a peak of 1,313 people killed in 1966 to 397 in 2008. As a first step to “Vision Zero”, i.e., no fatalities in road traffic, the Swedish government has launched a programme aiming to reduce the number of people killed to a maximum of 220 in the year 2020. Besides a variety of road safety measures, the programme includes monitoring 13 traffic safety indicators. One of these indicators is the Vehicle Speed Index on national roads.

Speed trends have been monitored in Denmark since 2002 in the Hastighedsbarometer, which every month presents average vehicle speeds from 5–8 measurement sites on a number of road types. A speed index is currently under development in Norway (Sakshaug 2009). In Sweden, a large survey of vehicle speeds on national roads and in urban areas was conducted annually between 1996 and 2004 (see Isaksson 1997; 2003) and (Forsman et al. 2003). This survey was based on a probability sample of measurement sites (1,600 sites on national roads and 800 in urban areas) and produced estimates of average vehicle speeds and several other variables. However, the survey was discontinued
in 2005 because of a budget cut. Instead, a speed index was developed based on a measurement system originally designed for estimating the change of total vehicle mileage on national roads. This system has been operating at 83 measurement sites since 1981, where traffic count is the primary variable and speed data is collected as a side-product. Although the Vehicle Speed Index is very approximate, it is still useful for various purposes.

The design of the measurement system is presented in Section 2. In Section 3 we discuss how we use the speed data for the index calculations. The index formulas are presented in Section 4, the resulting index series in Section 5 and the estimation of confidence bounds in Section 6. Section 7 deals with two aspects of the interpretation of the index and Section 8 contains some concluding remarks.

2. The Speed Data: Design of the Change of Total Vehicle Mileage Survey

Here we will briefly present the design of the Change of Total Vehicle Mileage (CTVM) survey, which provides the data used for the speed index.

In CTVM, the national road network is divided into road links, in such a way that the traffic flow is constant within each road link. The links are stratified according to seven regions and four road types. The total sample size is 83, and the links are Neyman-allocated (following the principles of Neyman 1934) over the 28 strata. Within each stratum, the links are further stratified according to traffic type. The number of substrata is equal to the sample size and the total vehicle mileage in each substratum is relatively similar to the totals in the other substrata. In this way, the sample size equals one in each substratum, and links are drawn with a probability proportional to size (pps), with probabilities proportional to the total vehicle mileage of the links. The data collection is conducted at permanent count sites.

Speed data is saved over 1-hour periods as the mean speed of individual vehicles. In addition, the speed of individual vehicles is saved in speed classes of 5 km/h over 24-hour periods, which makes it possible to calculate the percentage of drivers exceeding the speed limits.

3. The Use of CTVM Data for Monitoring Speed Development

When measuring vehicle speed, the road network being studied can be defined in different ways. The researcher may choose to collect data regarding the whole road network or limit the data collection to the parts of the roads where the drivers can choose their speed without being inhibited by the road curvature. The latter case usually limits the road network being studied to straight parts of the roads. However, no consideration was taken to speed measuring for CTVM when the survey was designed. The sites were first selected randomly on the sampled links. However, if they were located on a curve or close to a junction, they were moved to the closest suitable location with the same traffic flow as

A location is “suitable” if the equipment can be installed safely and at a reasonably low cost. This normally means straight roads.
the first selected location. Therefore the road network being studied can be roughly defined as “straight roads.”

The vehicle speed index allows the SNRA to monitor on a monthly basis the effects of general safety measures for vehicle speeds, including fining or information campaigns. Local measures, such as setting up road safety cameras or changing the speed limit, are not reflected in the index. This is because these measures have a profound impact on the vehicle’s speeds and measurement sites close to new cameras and new speed signs may be under- or overrepresented in the small sample of 83 sites. Measurement sites close to new cameras and new speed signs have therefore been taken out of the sample for at least one year.

Road-works close to the measurement sites may dramatically lower vehicle speeds. Information about road-works is not available when the index is calculated. However, in the data editing process, periods of abnormal speeds can be identified. In these cases, a number of days or the whole month (if the number of days with abnormal speeds is greater than 10) can be taken out of the index calculations.

The sample of 83 CTVM measurement sites is a two-stage probability sample where road links are primary sampling units (psu) and one measurement site is randomly selected within each psu. However, for the speed index, we do not use the inclusion probabilities, since the calculation of a speed index is not a strict estimation procedure with a well-defined population parameter. 5

Each of the 83 measurement sites is measured in two directions. Each direction is regarded as a separate sampling unit. One reason for this is that the vehicle speeds in the two directions may differ depending on, for example, the curvature. Another reason for separating the two directions is to minimise the missing data; if one direction is missing, data from the other may still be used. The correlation between the two directions is overlooked. The total sample size for the speed index is therefore 166.

Since the Change of Total Vehicle Mileage survey was not designed for estimating speed, there was no stratification of the measurement sites on speed limits. Instead, post-stratification on speed limits was performed in the following way (Table 1).

The stratum weights are the percentages of total vehicle mileage. The weights are updated annually.

Table 1. Post-stratification of the measurement sites and the secondary sampling units

<table>
<thead>
<tr>
<th>Speed limit</th>
<th>Road width</th>
<th>Number of measurement sites</th>
<th>Number of sampling units (two directions on each measurement site)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 km/h</td>
<td></td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>70 km/h</td>
<td></td>
<td>21</td>
<td>42</td>
</tr>
<tr>
<td>90 km/h</td>
<td>&lt; 8 metres</td>
<td>15</td>
<td>30</td>
</tr>
<tr>
<td>90 km/h</td>
<td>&gt; 8 metres</td>
<td>22</td>
<td>44</td>
</tr>
<tr>
<td>110 km/h + highways</td>
<td></td>
<td>18</td>
<td>36</td>
</tr>
</tbody>
</table>

5 A population parameter may be vaguely defined as “the vehicle speed development on (straight) national roads.”
4. The Speed Index

4.1. Variables of Study

Three variables are monitored, each with a separate index: average speed,\(^6\) percent drivers exceeding speed limit,\(^7\) and drivers exceeding the speed limit by more than 5 km/h.

4.2. Index Parameters of Interest

Each of the variables of study is monitored with two parameters:

i) Annual development per calendar month. The development of the average speed (or the number of drivers exceeding the speed limit) between month \(t\) and the same month one year before, two years before, and so on. The index is constructed as a chain index where index links representing 12-month periods are multiplied. Data from the year 1996 and later is available. Since the links are annual, no seasonal adjustment is needed. This index parameter describes the long-term development of the variables.

ii) Development per month. The development per month of the average speed (or the number of drivers exceeding the speed limit). This index is constructed as a chain index where links representing one-month periods are multiplied. The index is seasonally adjusted and smoothed with double exponential smoothing. Data from 2004 and later is available. This index parameter is primarily useful for studying the short-term effects of measures taken to lower vehicle speeds.

Here we restrict our attention to parameter (i), which is regarded as being the most important.

4.3. Index Formula for Average Speed

4.3.1. Annual Index Link Within a Post-stratum

An index link for the development of the average speed between a given month \(t\) and the same month one year earlier in post-stratum \(h\), is calculated as the geometric mean of the speed developments in sample units \(i\).

\[
L^j_h = \prod_{i=1}^{n^j_h} \left( \frac{\bar{x}^j_{hi}}{\bar{x}_{hi}} \right)^{w^j_{hi}}
\]

(1)

where:

- \(\bar{x}^j_{hi}\) = average speed in site/direction \(i\), month \(t\) in year \(j\), stratum \(h\). The base year is \(j = 0\).
- \(n^j_h\) = number of common\(^8\) sites/directions year \(j\) and \(j - 1\) for month \(t\), stratum \(h\).

\(^6\)Average speed is based on all vehicles.
\(^7\)The percentage of drivers exceeding the speed limit is based on only private cars without a trailer, since other vehicles may have speed limits that differ from the speed signs.
\(^8\)The common sites are those sites measured in both months.
4.3.2. Annual Index Link Over $H$ Post-strata

For an annual link over all $H$ post-strata for speed development between year $j - 1$ and $j$ for month $t$, the different stratum indices in (1) are weighted. The index formula is the geometric mean:

$$L^{j,t} = \prod_{h=1}^{H} \left( L^{j,t}_{h} \right)^{W^{j,t}_{h}}$$

where the weight $W^{j,t}_{h}$ is the percentage vehicle mileage in post-stratum $h$, which can change over a long period of time.

Sometimes it is interesting to have only one index link for each year and this index can be defined as the geometric mean over the twelve months

$$L^{j} = \prod_{t=1}^{12} \left( L^{j,t} \right)^{1/12}$$

4.3.3. The Chain Index Over $k$ Years

The vehicle speed index $I^{k,t}$ for a certain calendar month $t$ over $k$ years is constructed by multiplying the annual links.

$$I^{k,t} = \prod_{j=1}^{k} L^{j,t}$$

The mean index over $k$ years is:

$$I^{k} = \prod_{j=1}^{k} L^{j}$$

4.4. Index Formula for the Percentage of Drivers Exceeding the Speed Limit

We denote with $y^{j,t}_{hi}$ the percentage of drivers exceeding the speed limit during month $t$ in year $j$ at sampling unit $i$, post-stratum $h$. An annual link within stratum $h$ is defined as:

$$L^{j,t}_{h} = \frac{\sum_{i=1}^{n_{h}^{j,t}} w^{j,t}_{hi} y^{j,t}_{hi}}{\sum_{i=1}^{n_{h}^{j-1,t}} w^{j-1,t-1}_{hi} y^{j-1,t-1}_{hi}}$$

Formula (6) is motivated by the fact that $y^{j-1,t}_{hi}$ can be very small and sometimes equal to zero. We therefore cannot use the natural geometric mean as in formula (1).\(^9\) The rest of the links and indices are then defined analogously to Formulas (2)–(5).

\(^9\) It is easy to find empirical evidence for the similarity between Formula (6) and the geometric mean of the $y^{j,t}_{hi}/y^{j-1,t-1}_{hi}$ as long as the $y^{j-1,t-1}_{hi}$ are substantially $> 0$. 
5. Presentation of the Index

The Vehicle Speed Index is presented monthly on the Swedish National Road Administration’s official website, see http://vv.se/Startsida-foretag/vagar/Vagnet/Trafikfloden-och-medelhastigheter/Hastighetsindex/

For each of the variables (average speed, the percentage of drivers exceeding the speed limit, and the percentage of drivers exceeding the speed limit by more than 5 km/h) the index curve is displayed in a figure at the website. We illustrate this in Figures 1–3 by showing the indices for the month of October between 1996 and 2009. Typically, the three indices show a similar pattern, although the scales differ.

Fig. 1. Annual development for October 1996–2009. Average speed

Fig. 2. Annual development for October 1996–2009. The percentage of speeding offenders
6. Confidence Limits

6.1. A Bootstrap Procedure

The variability of the chain index is calculated with a Bootstrap procedure (Efron and Tibshirani 1998). In the Bootstrap procedure, 500 pseudosamples for each link were drawn with simple random sampling with replacement from the sample for all months and years. Recall that the original sample of 83 sites was a stratified sample from 28 strata. If you follow the Bootstrap theory in strict terms, the resampling should be made from these strata with sample sizes equal to the actual stratum sample sizes. However, this was not possible due to practical circumstances. Instead the resampling utilised the post-strata with fixed stratum sizes, i.e., the sample size was \( n_{ij} \) in each post-stratum. The Bootstrap samples therefore miss the variability caused by the post-stratification procedure, but on the other hand contain variability emanating from possible differences between the 28 CTVM strata.

From Formula (4) the chain index after \( k \) years \( I_{k,t} = \prod_{j=1}^{k} I_{r}^{j,t} \) was calculated in each pseudo-sample, \( r \), for each possible \( k \) as well as the variances

\[
\text{Var}(I_{k,t}) = \frac{1}{499} \sum_{r=1}^{500} \left( I_{r}^{j,t} - \frac{1}{500} \sum_{r=1}^{500} I_{r}^{j,t} \right)^2 \quad k = 1, 2, \ldots, 12
\]

We have not used the common percentile method (Efron and Tibshirani 1998, p. 170) for constructing confidence intervals, but instead the simple formula \( I_{k,t} \pm 2\sqrt{\text{Var}(I_{k,t})} \). This is motivated by reasons of simplicity, but has some support in the empirical distribution of the pseudovalues, which are approximately normally distributed.

In Figures 4–5, the confidence limits are explored for \( k = 1, 2, \ldots, 12 \). The bold lines are the indices for average speed for June and February and the broken lines are the values of upper and lower confidence limits.
6.2. A Logarithmic Transformation

We noted above that the indices were approximately normally distributed in the pseudo-samples. Using a logarithmic transformation we receive a better approximation of the normal distribution, and all products in the calculations are transformed to sums. If we overlook the correlations between the links, the variance of a logarithmic index is the sum of the variances for the logarithmic links. This independence assumption considerably simplifies the bootstrapping method and the use of the results. Each logarithmic link will now be bootstrapped separately and the variance computed. With this assumption, each logarithmic index of interest can be simply computed and its variance is the sum of

---

Fig. 4. Confidence limits for average speed, June 1996–2008

Fig. 5. Confidence limits for average speed, February 1996–2009
the actual logarithmic link variances. The confidence interval for an index can now be computed as:

$$\exp \left( \sum_j \ln(L_j) \pm 2 \sqrt{\sum_j \text{Var}(\ln(L_j))} \right)$$

To find some empirical justification for the independence assumption, we compared the confidence intervals for the average speed index for the month of June computed by the two bootstrapping methods. As is shown in Figure 6, the results are very similar, and our conclusion is that the simple logarithmic method is adequate for our purposes.

7. Issues Regarding the Interpretation of the Index

7.1. The Effect of a General Measure: Increased Fines

The exclusion of measurement sites because of new cameras and new speed limits leads to an important limitation of the indices: the indices do not reflect the actual speed development on the roads but only the effects of general measures, such as increased fines, to reduce vehicle speeds. One opportunity to verify this is to look again at the October indices, Figures 1–3 above. By 1 October 2006, the fines for speeding had doubled in Sweden. Moreover, during the first week in October 2006, the police conducted many speed checks and were generally very visible on the roads all over the country. The October indices for 2006 fell, as was to be expected.
7.2. The Effect of Weather on Vehicle Speed

Vehicle speeds are affected by different weather conditions, especially during winter periods. Periods when roads are in a winter state will give lower mean speeds than periods when roads are free of snow and ice. Figures 7–8 illustrate the differences in average speed between summer months (April to September) and winter months (October to March). In the summer months the weather conditions are almost the same, and we notice very stable downward trends in each month. The winter figure is completely different. The fluctuations between years and between months are very high, and the probable explanation is the fluctuating weather conditions during the winter period.

Fig. 7. Annual development for summer months (April to September). Average speed

Fig. 8. Annual development for winter months (October to March). Average speed
8. Concluding Remarks

Traffic data is expensive to collect, which means that data often has to be used for multiple purposes. The Vehicle Speed Index uses speed data from sites originally designed for measuring traffic flows. In this paper, we have discussed various practical concerns when this data is used for measuring vehicle speed development.

An advantage with a rough, low-budget index is that it can be running continuously for monitoring speed development. In addition to the Vehicle Speed Index, the Road Administration is planning to conduct full-scale surveys of vehicle speed every four years.

Further development of the Vehicle Speed Index will include

- The development of an Index calculator on the website with the option to select a base year. The calculator can also serve to calculate confidence limits using the logarithmic approach described above.
- This type of index for summer months (April to September) would seem to be easily interpretable since weather fluctuations do not affect the vehicle speeds. Such an index, as well as its confidence limits, can easily be calculated by the methods discussed in the paper.

9. References


Sakshaug, K. (2009). Fartsindeks: Krav til tellepunkt som skal ingå i fartsindeksen. Internal memorandum Sintef, Trondheim. [In Norwegian]