Roar Norvik
Norwegian Public Roads Administration
Department of Technology
Centre for Road and Traffic Technology, Trondheim
roar.norvik@vegvesen.no

Processing traffic data for statistical use; The Norwegian methods

This paper considers traffic data in Norway. The first chapter presents how data are collected as a background for the following chapters. Chapter 2 gives a short overview of the system that processes and stores traffic data. The main focus of this paper is the methods that are used in the traffic data system. The basic curve method is a method to estimate yearly traffic volume at a count site, when traffic counts are available for only a limited period, from a few days up to a week. Another method that is to be described is how the traffic growth is calculated.

1 Introduction

A huge effort is used in the process of collecting and processing traffic data, with over 7500 counting stations in use. The counting stations are separated in three different levels, based upon how often they are used to register. The level 1 counting stations are collecting data continuously. Figure 1 shows the geographic placement of these stations. There are around 250 of them. Level 2 are stations which only register data 4 weeks every fourth year. Level 3 collects data one week every sixth year.

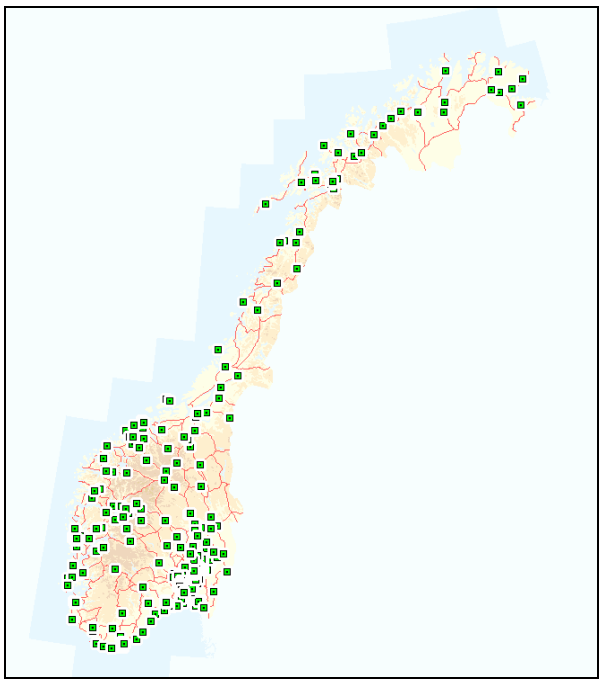

Figure 1 Level 1 count stations in Norway

The equipment used to collect data is based on traditional inductive loops cut down in the road (figure 2). The loops are connected to a computer. The computer has a modem that we use call to collect the counting-files. The files are then stored in the database described in the next chapter.

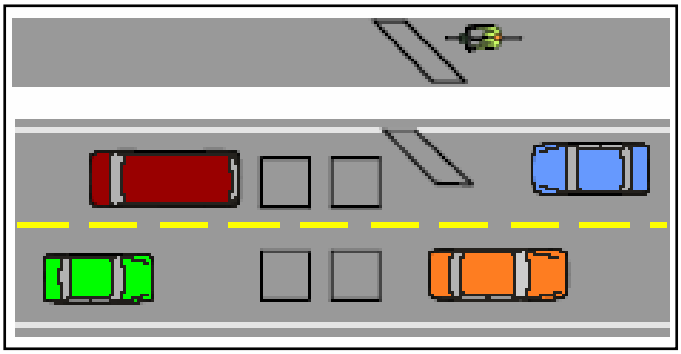

Figure 2 Inductive loops at a count station

For every passing vehicle a large amount of data is registered, but the most important data are length, direction and speed. We separate the vehicles into 5 length classes: 0 – 5.5 meters, 5.6 – 7.5 meters, 7.6 – 12.4 meters, 12.5 – 15.9 meters and longer than 16 meters.

2 The traffic data system

The official data system for processing traffic data is NorTraf. The system consists of a PC-application and an Oracle-database on server. The application consists of 7 modules that execute the following:

KR2000
A module for input, control and correction of data files. The user is allowed to set the controls that are necessary to verify the quality of the data.

Calculation
A module that uses the basic curve method to calculate several traffic volume parameters such as annual average daily traffic (AADT). The method is described in chapter 3.

Tellius
A program to add new and edit existing counting stations. It is also called the count station registry.

TrafKurv
A module to present and print different variation curves.

NorTrafAdm
A report-module to produce traffic data reports. Official traffic data statistics such as the traffic index are to be reported from this module.

ÅDT2000
A module to assign annual average daily traffic to the road network.

Map-module (*under development*)
Used to establish counting stations in Tellius and to present point- and network data in a map.

3 The basic curve method[1]

When estimating traffic volumes on a yearly basis the variation curve method has traditionally been used. The curves were static and seldom updated. That method was used to calculate traffic data in the Traffic database. When developing the new traffic data system NorTraf, the basic curve method was implemented. The method provides an estimate of the yearly traffic volume at a count site, when traffic counts are available for only a limited part of the year, from a few days up to a week. The results are several traffic volume parameters such as average annual daily traffic (AADT), average annual workday traffic (YDT), average annual weekend traffic (HDT), average annual summer vacation traffic (SDT) and average annual July traffic (JDT). All the parameters are presented with their individual uncertainty. This chapter presents a simplified description of the method.

3.1 Formula

The basic curve method can be described as a function b_t of hour t in the year given by

$$b_t = aa_t * ud_t * tr_t * s_t$$

where

aa_t	=	the yearly variation
ud_t	=	weekly and daily variation
tr_t	=	trend, a long term increase or decrease
s_t	=	special days (Easter, Christmas and other public holidays)

The amount of vehicles per hour in hour t is estimated by

$$a_t = c * b_t$$

where c can be found from the available counted data, so that the adjusted curve passes through the counted data. I.e. that c has the same level as the counted data

The basic curve method is in many ways similar to the variation curve method. One difference is that the new method allows for a trend over time, and special days like Easter, Christmas and others influence the result. Another difference that is important is this:

The function b_t can also be written as

$$b_t = b_{1t} * k_t$$

where

- b_{1t} is a predefined function, independent of the counted data, called the first basic curve
- k_t is a correction factor dependent of the counted data

If there is little counted data available $k_t = 1$, i.e. no correction. If there are more counted data available k_t is a function of the data. The basic curve method has a flexibility that is adaptive to the amount of data.

The basic curve method has been calibrated on data registered in about 100 level 1 counting stations.

3.2 Curves

Two sets of basic curves have been developed. One set to be used in larger cities and one set to be used in smaller cities and rural areas. The main difference on a yearly basis in these curves is shown in the figures below. The City-curve has less traffic in the easter- and summervacation. The Land-curve has more traffic in the same periodes.

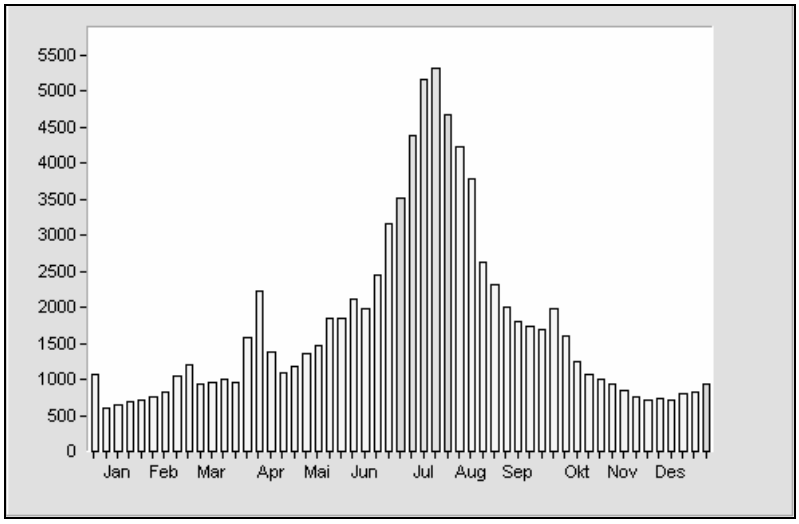

Figure 3 A typically curve for smaller cities and rural areas.

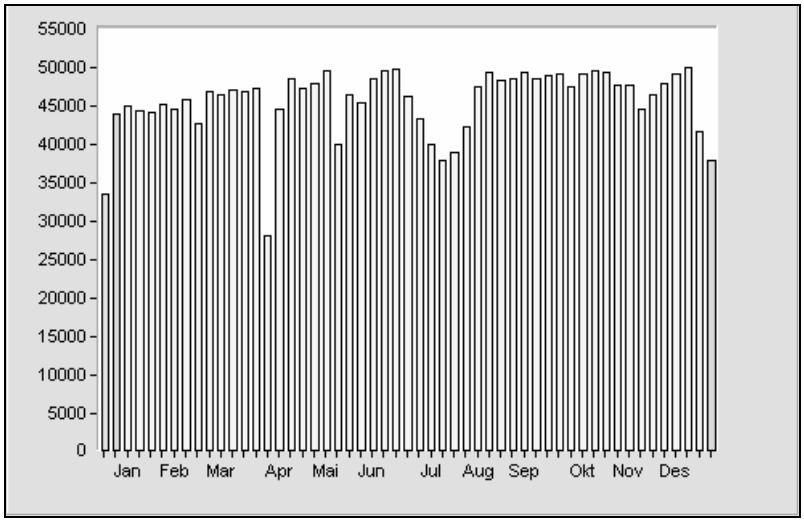

Figure 4 A typically curve for larger cities.

For each group there have been developed basic curves for each length class. The figure below shows a decomposed basic curve in a larger city for length class 0 – 5.5 meters.

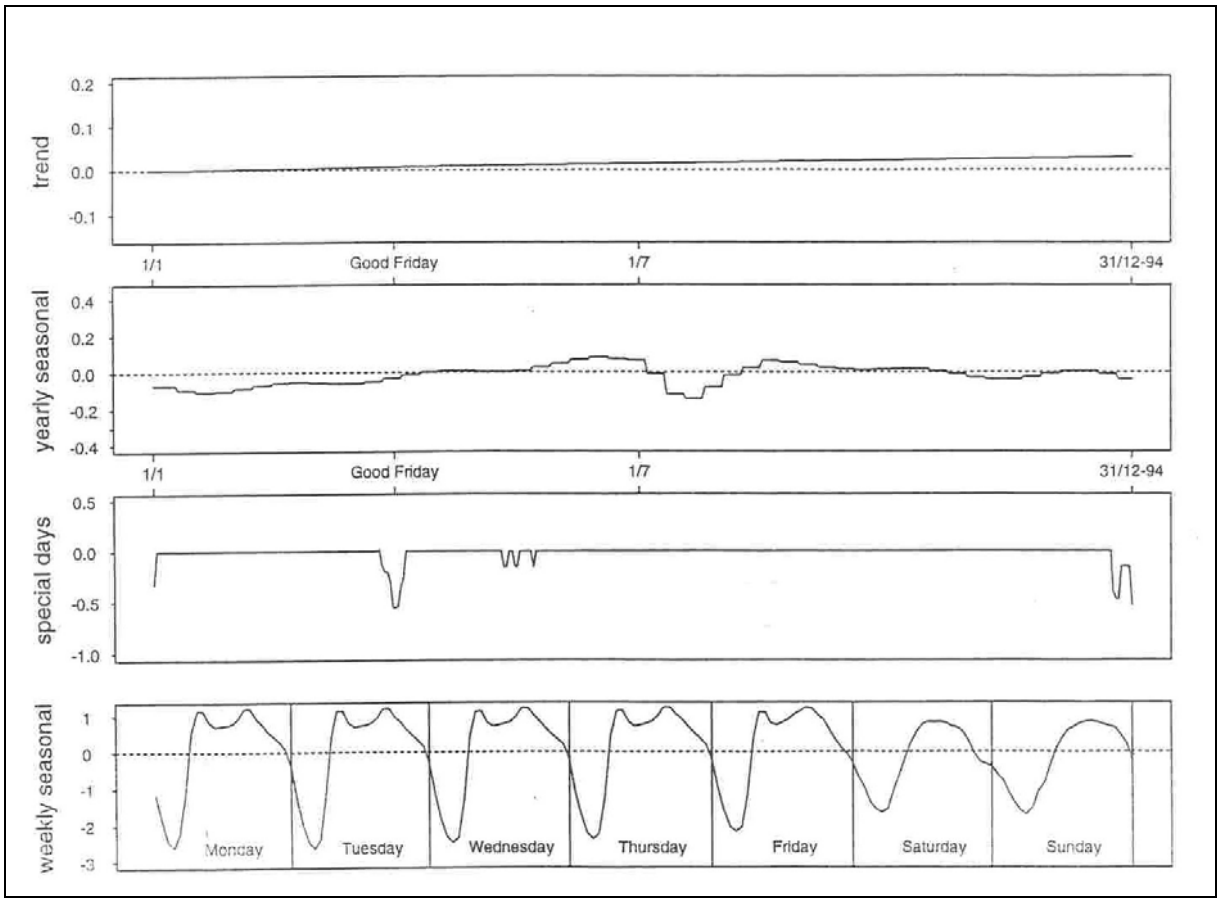

Figure 5 First decomposed basic curve for length class 0 - 5.5 meters.[i]

If we take a look at the seasonal variation curve it is a typical sign that the traffic volume is lower in the summer. The daily variation curve shows a rush in the morning and in the afternoon.

The critic against this method has been that it has only 2 sets of curves. What about roads with extreme summer traffic or roads that are closed during the winter? Today this is solved by counting several periods during the year. In the future the intension is to develop more curves.

3.3 Uncertainty[ii]

The uncertainty of an AADT-estimate depends on when and for how long the traffic is counted. In general the uncertainty will decrease when the counting period becomes longer, but it also depends on what time of day or week the counts are done. The uncertainty will also typically vary as a function of the traffic volume.

The figure below shows how the relative standard error in the AADT-estimate decreases towards 0 when the counted period increases from 1 to 52 weeks for length class 0 -5.5 meters. We can also see that the uncertainty decreases at higher traffic volumes.

[i] The curves in the figure are based upon data from the year 1994. The curves that are used in the application are based upon data from the year 2000. This figure shows the principle in the method.

[ii] This paper doesn't focus much on uncertainty, but it is mentioned because it is a part of the basic curve method. No formulas will be presented.

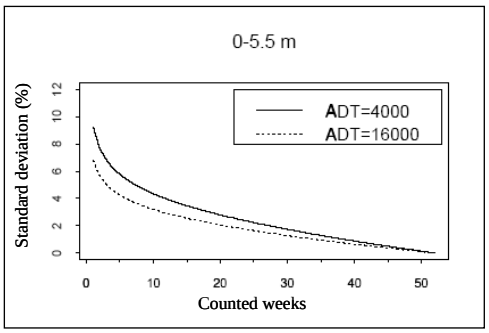

Figure 6 Relative standard error in AADT-estimate for length class 0 - 5.5 meters

4 The traffic index[2]

The traffic index is the official Norwegian index for traffic growth. The index is a measure of the increase or decrease of the annual vehicle distance travelled (AVDT) within a specific region or for the whole country. The AVDT is defined as the total amount of kilometres travelled by vehicles at a given period inside a geographic area:

AVDT=AADT * lenght[iii] of road section

The index provides us with a number on the traffic growth on a monthly basis, a total so far this year and the last 12 months for single count stations, regions, counties and total for the country. Today the index is calculated with permanent weights. These weights are more than 10 years old and derive from a calculation of the total AVDT of each county.

The index is now being revised. A new index that contains floating weights instead of permanent weights will be implemented in NorTraf. This chapter describes the implementation of the new index.

Only level 1 count stations[iv] is used in the calculation of the index.

4.1 Selection

The index calculation is about comparing data from one year, called the reference year, with the following year (the index year). It is important to choose the right selection.

For every hour, in every length class, in each lane, in the same time-interval we have to check if data exists both in the reference year and the index year. If that holds true, we bring the data from both years into the selection.

Example
For length class 0 – 5.5 meters, in direction R1, in lane 1, on May[5th] in 2003 and 2004, and for each hour there has to exist data in X1, X2… and Y1, Y2… . If that holds true we bring the sum X1-X24 and Y1-Y24 into the calculation.

Length class	< 5,5 m		Length class	< 5,5 m
Direction	R1		Direction	R1
Lane	1		Lane	1
Date	**05.05.2003**		Date	**05.05.2004**
Time			Time	
00:00	X1	**< - >**	00:00	Y1
01:00	X2	**< - >**	01:00	Y2
02:00	X3	**< - >**	02:00	Y3
03:00	X4	**< - >**	03:00	Y4
04:00	X5	**< - >**	04:00	Y5
05:00	X6	**< - >**	05:00	Y6
06:00	X7	**< - >**	06:00	Y7
07:00	X8	**< - >**	07:00	Y8

[iii] The length is set to 15 km for all counts and it is therefor left out of the calculation. It is our intension to find the length of all road section by analyzing the network and implementing this in the calculation of the index.
[iv] Data collected continously

08:00	X9	**< - >**	08:00	Y9
09:00	**X10**	**< - >**	09:00	**Y10**
10:00	X11	**< - >**	10:00	Y11
11:00	X12	**< - >**	11:00	Y12
12:00	X13	**< - >**	12:00	Y13
13:00	X14	**< - >**	13:00	Y14
14:00	X15	**< - >**	14:00	Y15
15:00	X16	**< - >**	15:00	Y16
16:00	X17	**< - >**	16:00	Y17
17:00	X18	**< - >**	17:00	Y18
18:00	X19	**< - >**	18:00	Y19
19:00	X20	**< - >**	19:00	Y20
20:00	X21	**< - >**	20:00	Y21
21:00	X22	**< - >**	21:00	Y22
22:00	X23	**< - >**	22:00	Y23
23:00	X24	**< - >**	23:00	Y24

Data can be excluded if special incidents occur that change the route choice for many vehicles. Examples can be road maintenance, traffic accidents etc. Bigger events (for instance the Winter Olympic Games in Lillehammer 1994) that increase or decrease the total traffic in an area for a longer period are not excluded, but have to be a part of the general growth.

4.2 Calculation of several indexes

4.2.1 Calculation data

At first we add the numbers for each count station, for each month, for each length class for both the reference year and the index year. (The process is here simplified; we have aggregated data for each lane and each direction). We then get a monthly sum of traffic (sum) and the amount of hours counted (count) (to find the average traffic per counted hour). For both the reference and the index year data is aggregated for every day (all), for workdays and for weekend-days. The examples are only shown for every day (all).

Example

Calculating the index for the sum of traffic for every day (all), both directions, for one length class and for one count station.

At first data is aggregated for a month for one count station for the reference and the index year.

$$sum\_all\_reference = \sum_{day=1}^{28/30/31} \sum_{hour=0}^{23} \sum_{lane=1}^{2/3/..} value(day, hour, lane)$$

$$sum\_all\_index = \sum_{day=1}^{28/30/31} \sum_{hour=0}^{23} \sum_{lane=1}^{2/3/..} value(day, hour, lane)$$

The amount of hours in the selection is also summed up:

$$count\_all = \sum_{day=1}^{28/30/31} \sum_{hour=0}^{23} \sum_{lane=1}^{2/3/..} 1$$

4.2.2 Index for one count station for one month

$$index\_month\_all = (\frac{sum\_all\_index}{sum\_all\_reference} - 1) * 100$$

4.2.3 Index for one count station, for the whole time-interval

The months have to be weighted

Index for all days, for each length class

$$index\_countstation\_all = ((\sum_{month=1}^{n} index\_month\_all(month) * weight(month)) - 1) * 100$$

where

n = the number of months calculated

Calculating the weights(month)
To calculate the weights, we use the average monthly traffic[v] calculated by the basic curve method in the calculation module in NorTraf.

$$weight(month) = \frac{monthly\_traffic(month)}{\sum_{i=1}^{n} monthly\_traffic(i)}$$

where:

month: the month in which we are calculating the weight
monthly_traffic(i): picked up from the database, for month i

Notice the interdependence: $\sum_{i=1}^{n} weight(i) = 1$

4.2.4 Index for all count stations in one county

The count stations are now being aggregated to one county.

The index for all days (all), for each length class:

$$index\_county\_all = ((\sum_{countstation=1}^{n} index\_countstation\_all(countstation) * weight(countstation)) - 1) * 100$$

[v] From the reference year.

where

n = the amount of count stations within a county

Weight(countstation)

$$sum(countstation) = \sum_{month=1}^{n} \left(\frac{sum\_all\_reference(countstation, month)}{count\_all(countstation, month)}\right)$$

$$weight(countstation) = \frac{sum(countstation)}{\sum_{i=1}^{m} sum(i)}$$

where

n = the number of months calculated
m = the amount of count stations
sum(countstation) = the relative AVDT in one count station

Notice the interdependence: $\sum_{i=1}^{m} weight(i) = 1$

4.2.5 Index for several/all counties

If you are interested in calculating the index for a multiple of counties such as a region, province or the whole country, we need to give weight to the counties too.

Index for all days, for each length class:

$$index\_counties\_all = \sum_{county=1}^{n} index\_county\_all(county) * weight(county)$$

where

n = the counties you are calculating for

Calculating the weights
To calculate these weights we use AVDT's for the reference year for each county. These numbers are calculated in another process but are stored in the database.

The weights become: $weight(county) = \frac{AVDT(county)}{\sum_{i=1}^{n} AVDT(i)}$

where

n = the counties you are calculating for

Notice the interdependence: $\sum_{i=1}^{n} weight(i) = 1$

5 Summing up

Implementing the basic curve method and the new calculation of the traffic index in the traffic data system NorTraf has resulted in more reliable traffic statistics in Norway.

The basic curve method is able to produce more precise AADT estimates than the variation curve method. The main reason is that the basic curve method allows a trend over time and is influenced by special days like Easter and Christmas. It also has a correction factor which is a function of the amount of data available. The basic curve method also estimates the uncertainty of an AADT-estimate in the shape of the relative standard error.

The new official traffic index has a better method for selecting data. Excluding special incidents from the data sets are now being implemented in case of road maintenance, traffic accidents etc. We can calculate the index at several levels. For example index for: one count station for one month, one count station for several months, several count stations in one county and an index for a multiple of counties such as a region or the whole country.

References

Aldrin, M. Beregning av trafikkvolum ved hjelp av basiskurvemetoden – En innføring SAMBA /05/00 (Calculation of traffic volumes by means of the basic curve method – An Introduction)

Aldrin, M. Traffic volume estimation from short period traffic counts. Norwegian Computing Centre, Oslo

Johansen, K. Vegtrafikkindeksen – metode og funksjonsbeskrivelse, Statens vegvesen Vegdirektoratet 2004 (The new traffic index – A description of the method and the functionality)

Vestues, T. Indeksberegning – dokumentasjon, VegInformatikk 2004 (The calculation of the index – The documentation)

[1] Developed by Magne Aldrin at the Norwegian Computing Centre
[2] Developed by Kjell Johansen, Norwegian Public Roads Administration