

## **Weigh-in-Motion – Categorising vehicles**

D.W. van Boxel, MSc.<sup>1</sup>  
Ing. R.A. van Lieshout, MSc.<sup>1</sup>  
R.A. van Doorn, MSc.<sup>2</sup>

<sup>1</sup> CQM bv, P.O. Box 414, 5600AK, Eindhoven, The Netherlands, tel. +31 40 27 57747, e-mail: info@cqm.nl

<sup>2</sup> Ministry of Transport, Public Works and Water Management, Road and Hydraulic Engineering Institute (DWW)  
P.O. Box 5044, 2600 GA, Delft, The Netherlands, tel. +31 15 251 8518, email: dwwmail@dww.rws.minvenw.nl

### **Abstract**

Overloaded vehicles cause great damage to the roads. The Weigh-in-Motion project of the Dutch Ministry of Transport intends to reduce the number of overloaded vehicles on the Dutch roads. In order to detect overloaded vehicles automatically, the axle loads of passing vehicles are compared to the legal limit. Since the legal limits differ for different types of vehicles or axles, it is critical that vehicles are classified correctly. Analysis of a sample showed that the existing algorithms could not, or incorrectly classified 17% to 30% of the vehicles. CQM developed an innovative classification method with an error percentage of merely 2%. This paper describes this method.

## **Introduction**

In 2001 the Road and Hydraulic Engineering Institute (DWW) of the Dutch Ministry of Transport, Public Works and Water Management built six Weigh-in-Motion (WIM) systems on the Dutch motorway network. These systems are used for pre-selection of overloaded vehicles, collecting data for overload prevention actions and road management statistics.

To ensure the functioning of the WIM system it is essential that axle loads of passing vehicles can be compared to the legal limits. These legal limits differ depending on the type and configuration of the vehicle. It is therefore crucial that vehicles are classified correctly when passing a WIM location. In addition to law enforcement these data are used to analyse the use of the infrastructure, which only increases the need for a correct classification.

Currently there are two classification algorithms in use on the Dutch motorway network. They both use axle separation. The original algorithm, which is referred to as Euro13, was supplied with the Marksman roadside equipment. The second, which is currently in use, is an adapted version of the Euro13 algorithm, this algorithm will be referred to as WIM-NL. In addition to these two algorithms there is a third algorithm, which was developed and is being used by the Dutch province of Zeeland. This algorithm has yielded results similar to the other algorithms.

## Data

To rate the performance of the existing algorithms and to test any new algorithm, CQM was provided a dataset containing 1691 vehicles. The data comprises pictures (see Figure 1) of all vehicles as well as the speed of the vehicle, axle count, axle loads, and axle separations. All these vehicles were then classified by eye to obtain the true class, as well as being classified using each of the three algorithms.



**Figure 1:** Sample photograph

Prior to any analysis, the data was split into two. This was done to assure that a new algorithm would not be favoured merely by the fact that it was being evaluated using the same data used to tune it. The first half can thus be viewed as a training set, the second as a verification set.

## Existing Algorithms

### Vehicle classes

There are a large number of possible vehicle class definitions. The standard class definitions used by the Marksman roadside system are referred to as the Euro13 classes. These classes distinguish between different types of vehicles (e.g. a Heavy Goods Vehicle or Bus), different configurations (with or without trailer) and different number of axles. A Euro13 class may be an aggregation of several of these. DWW has defined its own classes. These classes do not aggregate and are therefore more accurate. Each part of the vehicle is identified by a letter followed by one or more integers. Each integer indicates the number of axles in an axle group, where the first integer refers the first axle group etc. 'B<sub>12</sub>' for example is used to indicate a bus with a single axle at the front and an axle group of two at the rear. For a comprehensive definition see Van Doorn (2000). To compare the performance of the different algorithms a common set of vehicle classes should be used. For this purpose the Euro13 vehicle classification table is used.

## Euro13

The Euro13 algorithm is included as standard in the Marksman roadside system. It classifies vehicles solely on their axle separations. The algorithm is based on a flow diagram approach. If a set of axle separation conditions is met, it is classified as type *a*, if they are not met, the conditions for type *b* are checked etc. If none of the conditions are met the vehicle is classified as *other*. Since the boundaries set in each of the conditions are firm, this approach is particularly sensitive to measurement errors.

Table 1 lists the true vehicle class against the classification as computed by the Euro13 algorithm for the training set. If all vehicles had been classified correctly by Euro13 all entries should lie on the diagonal. All off-diagonal elements should thus be viewed as being erroneous. In the training set 250 vehicles are classified incorrectly, which corresponds to 29.7%. For the verification set the corresponding figures are 269 and 32% respectively. Two main problems are that the algorithm classifies many two axle heavy goods vehicles (HGV's), category 2, as busses, category 12. In addition it classifies many vehicles as *other*, category 13.

**Table 1:** Training set, 29.7% classified incorrectly

TRUE	euro13													Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	
1	27	40			16		6					1	4	94
2		76										110	7	193
3			29									3		32
4				12					1				5	18
5					51		1						3	55
6					2	13			9				2	26
7					1		50						3	54
8					6			124						130
9									158				18	176
10									2					2
11										15			4	19
12		1	7									23		31
13													13	13
Total	27	117	36	12	76	13	56	125	158	12	15	137	59	843

## WIM-NL

The WIM-NL algorithm is the modified version of the Euro13 algorithm which is currently in use. The WIM-NL algorithm differs from the Euro13 in that the axle separation bounds used as criteria for the different categories have been adjusted. In addition category 12 (busses) is no longer allocated to any two-axle vehicle. Table 2 show the performance of the algorithm on the training set. On this set the algorithm incorrectly classifies 139 vehicles, which corresponds to 16.5%. On the verification dataset these figures are 150 and 17.8% respectively.

**Table 2:** Training set, 16.5% classified incorrectly

TRUE	WIM-NL													Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	
1	72	1			14		6						1	94
2		193												193
3			29									3		32
4				12				1					5	18
5					24			1				1	29	55
6						10				14			2	26
7					1		52						1	54
8					16			113					1	130
9									169				7	176
10										2				2
11											15		4	19
12		24	7											31
13													13	13
Total	72	218	36	12	55	10	58	115	169	16	15	4	63	843

**In Sum**

Table 3 contains a summary of the error percentages for each of the two evaluated algorithms. Modifying the current algorithms by putting weight restrictions on each of the possible classifications as described in Table 6 would allow them to better detect cars and vans (class 1). Any vehicle below 3,500 kg would immediately be classified as a car or a van. Any tractor trailer combination with a total weight less than 7,000 would also be classified as a car or a van. The resulting performance of the modified algorithms are shown in Table 4.

**Table 3:** Summary

	Training set	Verification set	Average
Euro13	29.7%	32.0%	30.9%
WIM-NL	16.5%	17.8%	17.2%

**Table 4:** Summary of modified results

	Training set	Verification set	Average
Euro13	21.6%	23.8%	22.7%
WIM-NL	14.0%	15.7%	14.9%

**Classification and clustering**

Splitting a set of objects into two or more groups of more or less 'similar' objects is an often performed activity. Degree of similarity between objects is obtained by measuring how close objects are, based on several measured characteristics of these objects. Some applications of grouping may be found enlightening. The grouping of industries based on published performance indicators is a good example. The groups or clusters of equal industries found can be used as marketing input. Each cluster may have to be approached differently. A second example is an insurance problem. Their question was to find out whether clients can be clustered according to their risk of having a serious accident. Premium rates should be in line with this risk making premium for low-risk drivers lower than for high-risk drivers. Typical for these kind of problems, called *clustering problems*, is the fact that the number of clusters is not known beforehand.

The second group of problems, referred to as *classification problems*, are problems with an a priori known number of clusters. A typical example occurs when a hospital wants to diagnose a disease from external symptoms without hospitalisation. So treatment can start at an early stage. As a second example think of an existing but expensive measurement method to identify the class of an object. Finding discriminators of a cheaper kind to predict the result of the test as would have been obtained with the expensive measurement method is a typical classification exercise. Grouping of HGV's can be seen as classification problem similar to the second example. By using EURO13 the number of groups is exactly known.

### What classification method to use?

Basic statistical classification problems are encountered when an object can be a sample of two or more different populations. Statistical populations arise when there is some kind of (natural) variation between the objects. As an example consider the two populations defined by the heights of all Dutch men on the one hand and the population of all Dutch women on the other hand. Given a new measured height  $H$ , 1.78 m say, from which population is this height most likely drawn? The real difficulty arises due to the fact that the two populations are overlapping. If not, assigning a height to the correct population can be done with a 100% certainty. A discrimination test looks like,

$$\begin{array}{ll} \text{Assign } H \text{ to the population of men} & \text{if } d(H) > 0 \Leftrightarrow H > d^{-1}(0) = k \\ \text{Assign } H \text{ to the population of women} & \text{if } d(H) \leq 0 \Leftrightarrow H \leq d^{-1}(0) = k \end{array}$$

Where  $d(\cdot)$  is a function showing, on statistical grounds, a maximum amount of discrimination. It can also be shown that the function  $d(\cdot)$  measures the weighted distance between the measurement and the mean of the distributions [Kendall et al, 1983].

The logical equivalence of this for HGV classification is easy. The roadside system generates a fingerprint for a given car. The vector,  $\mathbf{U}$ , consisting of the length of the truck plus all the measured distances between axes can serve as an example,

$$\mathbf{U} = (L, d_{1,2}, \dots, d_{n-1,n})$$

Where  $d_{n-1,n}$  will be set to missing if  $n > N$ ,  $N$  being the number of axes of the truck under study. Building a function  $F$  that maps  $\mathbf{U}$  onto the set of EURO-13 classes ( $F: \mathbf{U} \rightarrow \Omega$ ) is the issue of our classification activity.

## CQM Approach

Using standard statistical techniques to classify vehicles did not yield very intuitive results. We developed an intermediate station that does not exist in the standard statistical classification theory, the transformation of the vector  $\mathbf{U}$  into an *interpretable* one-dimensional string  $S$ . A string  $S$  can be drawn based on the length of the vehicle and the distances between axles. As an example assume that the string  $\mathbf{U} = (700,50,200,50)$  meaning that the length of the truck  $L=700\text{cm}$ , the distances between two successive axes are 50, 200 and 50 respectively.

Figure 2 shows the transformation of this data.

$$\mathbf{U} = (700,50,200,50) \leftrightarrow \text{"-o-o-----o-o-----"}$$

**Figure 2:** Transforming data into a string

Each character in the vehicle signature string represents a length of 50 cm. Given this method of representing a vehicle with a string, the truck in Figure 3 can best be represented by string 1. In practice however it will be impossible to come to this string since it would require data on the overhang at the front of the vehicle. Data that cannot be retrieved from the roadside system at this time. A choice was made to start each string representing a vehicle with a single dash, followed by



- 1) ---o-----o---
- 2) -o-----o---

**Figure 3: Conversion of vehicle to string**

the first axle, the following dashes and O's representing the axles are given by the axle separation information. If the total length of this string corresponds to a total vehicle length less than the measured value, dashes are added to make up for this difference. Any overhang at the front of the vehicle would thus be added to the overhang at the back in the string representation. In the case of the example shown in Figure 3, this would result in the second string.

### Base strings

This string representing a passing vehicle can be compared with a set of *base strings* one for each category that is taken into account. The distance between the measured string and each base string can be calculated. The vehicle will be assigned to the category for which the distance between measured string and base string is the smallest. The training set was used to define a base string for each of these categories. This list is shown in the second column of Table 5.

### **Distance between two strings**

Given a set number of operators, such as inserting a character, removing a character and transposing two characters, a string can be turned into any other string. The string representation of a passing vehicle can thus be transformed to any of the base strings. Counting the number of edit steps required for each transformation gives us the distance between two strings. This distance allows us to identify the base string and corresponding vehicle class which best matches the passing vehicle.

Since the letter 'O' refers to an axle, it does not make sense to add or remove this character, this operation is therefore not permitted when transforming the vehicle string to any of the base strings. As a result only those base strings are considered that represent a vehicle class with the same number of axles as the passing vehicle.

Converting one string into another using the operators described above, can be done in many different ways. The objective however, is to use the minimal number of permitted edit steps so that a fair comparison can be made between the different base strings. To achieve this a Dynamic Programming algorithm is used.

### **Refinement**

On the road there is no such thing as a standard truck, bus, or articulated lorry. Dimensions differ to some degree from vehicle to vehicle within the same class. There is a wide variation of trailers, whereas the trucks are relatively uniform. When matching a vehicle signature string to one of the base strings it therefore makes sense to make it relatively 'cheap' to shorten a trailer in order to achieve a match, and make it 'expensive' to alter the length of the truck.

For this purpose the characters '~' and '\*' are introduced into the base string. When comparing the vehicle strings to each of the base strings these characters can both be read as '-'. Only when one of these characters needs to be removed or inserted to achieve a match will there be a difference. Inserting or removing the character '\*' is the most expensive, since this corresponds to altering the length of a truck; the character '-' can be removed or inserted against standard costs, '~' is used in the trailer section of the base string and can be considered cheap to remove or insert.

**Table 5:** Base strings

Vehicle class	Base string	Refined base string
V <sub>11</sub>	-○-----○-	-○-----○-
V <sub>21</sub>	-○-○-----○-	-○-○-----○-
V <sub>12</sub>	-○-----○○-	-○-----○○-
V <sub>22</sub>	-○-○-----○○-	-○-○-----○○-
V <sub>13</sub>	-○---○-○-○---	-○---○-○-○---
V <sub>11</sub> A <sub>11</sub>	-○-----○-----○-----○-	-○*****○-----○-----○-
V <sub>11</sub> A <sub>12</sub>	-○-----○-----○-----○○-	-○*****○-----○-----○○-
V <sub>11</sub> A <sub>1</sub>	-○-----○-----○-----○-	-○*****○-----○-----○-
V <sub>11</sub> A <sub>2</sub>	-○-----○-----○○-----	-○*****○-----○○-----
V <sub>12</sub> A <sub>11</sub>	-○-----○○-----○-----○-	-○*****○○-----○-----○-
V <sub>21</sub> A <sub>11</sub>	-○-○-----○-----○-----○-	-○-○*****○-----○-----○-
V <sub>12</sub> A <sub>1</sub>	-○-----○-----○-----○-	-○*****○-----○-----○-
V <sub>12</sub> A <sub>2</sub>	-○-----○○-----○-○-	-○*****○○-----○-○-
V <sub>12</sub> A <sub>12</sub>	-○-----○○-----○-----○○-	-○*****○○-----○-----○○-
T <sub>11</sub> O <sub>1</sub>	-○-----○-----○-----	-○*****○-----~::~~::~○-----
T <sub>11</sub> O <sub>2</sub>	-○-----○-----○○-----	-○*****○-----~::~~::~○○-----
T <sub>11</sub> O <sub>3</sub>	-○-----○-----○○○-----	-○*****○-----~::~~::~○○○-----
T <sub>12</sub> O <sub>1</sub>	-○---○○-----○-----	-○***○○-----~::~~::~○-----
T <sub>12</sub> O <sub>2</sub>	-○---○○-----○○-----	-○***○○-----~::~~::~○○-----
T <sub>12</sub> O <sub>3</sub>	-○---○○-----○○○-----	-○***○○-----~::~~::~○○○-----
B <sub>11</sub>	-○-----○-----	-○-----○*****-
B <sub>12</sub>	-○-----○○-----	-○-----○○*****-

**Post processing step 1**

It is possible that a vehicle which is to be classified is not in fact a bus or a heavy goods vehicle but a van or a car, with or without a trailer. To filter out these vehicles each category is assigned a minimum weight. Any vehicle below 3500 kg is assigned category *car*, any vehicle with three or more axles which weighs less than 7000 kg is also categorised as *car*, since these vehicles generally correspond to vans or cars with trailers. In addition each category is assigned a minimum weight, if this minimum weight is not met for a particular category, this category will be ruled out. The minimum weight of a tractor trailer combination is of course higher than that of a regular two-axle HGV. A vehicle which does not meet any minimum weights listed in Table 6 is assigned to Euro13 category 1.

**Table 6:** Minimal weights

Euro13 category	Weight
2	3,500
3	7,000
4	7,000
5	7,000
6	7,000
7	7,000
8	7,000
9	7,000
10	10,000
11	10,000
12	3,500

## Post processing step 2

The law prescribes when two adjacent axles can be seen as a pair and when they cannot. The new algorithm does not take such legal restrictions into account, a second post processing is therefore required. This step would take the classification as produced by the new DWW algorithm, from this it can easily determine the number of axles for both the HGV and the trailer. The procedure will then return to the original data and check it against the law. It will look at axle one and two, if the separation is less than the maximum prescribed by law for it to be considered a pair, the two will be said to be a pair etc. This procedure will be done for both the HGV and the trailer.

## Calibration

The construction of the base string can be viewed as a calibration of the algorithm. The strings listed in Table 5 were created based on a Dutch dataset. It is likely that for this approach to be used in other countries, with possibly different vehicle configurations, a re-calibration is called for. Re-calibration would also be required if new vehicle types are introduced, this however would apply to all classification methods.

## Results

Tables Table 7 and Table 8 show the performance of the new algorithm. The training set was used to construct the algorithm and could thus be viewed as bias. No automatic tuning of the algorithm has been done, it was therefore not expected that the algorithm would perform significantly worse on any other dataset. This expectation was reinforced by the results on the verification set. The new DWW algorithm classified 15 vehicles of the training set incorrectly, which corresponds to 1.8%. The verification set resulted in 20 incorrectly classified vehicles, corresponding to 2.4%.

**Table 7:** Training set, 1.8% classified incorrectly

TRUE	DWW													Total	
	1	2	3	4	5	6	7	8	9	10	11	12	13		
1	93											1			94
2		188											5		193
3			32												32
4				18											18
5					51			3	1						55
6						26									26
7			2				52								54
8				1					129						130
9										176					176
10											2				2
11												19			19
12		1	1										29		31
13														13	13
Total	93	189	35	19	51	26	52	132	177	2	19	35	13		843

**Table 8:** Verification set, 2.4% classified incorrectly

TRUE	DWW													Total	
	1	2	3	4	5	6	7	8	9	10	11	12	13		
1	94														94
2	1	206										3			210
3			25									1	1		27
4				12											12
5					61			1	4						66
6						37									37
7			1				48								49
8				3				99							102
9									186						186
10										1					1
11											11				11
12		4			1								35		40
13														7	7
Total	95	210	26	15	62	37	48	100	190	1	11	39	8		842

Combining the results achieved with the existing algorithm and those yielded by the new approach are given in Table 9. The results achieved through correcting the existing algorithms to better recognise cars and vans are given in Table 10.

**Table 9:** Results summary

	Training set	Verification set	Average
Euro13	29.7%	32.0%	30.9%
WIM-NL	16.5%	17.8%	17.2%
DWW	1.8%	2.4%	2.1%

**Table 10:** Corrected results summary

	Training set	Verification set	Average
Euro13	21.6%	23.8%	22.7%
WIM-NL	14.0%	15.7%	14.9%
DWW	1.8%	2.4%	2.1%

In addition to the significantly better results yielded by the new algorithm, the DWW algorithm is also extremely flexible. It is easy to add additional vehicle classes if new vehicle configurations are taken into use at some time in the future.

## Conclusions

The algorithms currently available have been shown to be inaccurate when used to classify vehicles passing the WIM locations. These algorithms could in theory be tuned to improve their performance on the dataset provided. No guarantee could however be given that such changes would yield improved performance of the live system. In addition the existing algorithms will always be sensitive to measurement error.

The new DWW algorithm is an innovative approach that has proved to be very effective when used to classify vehicles. The error percentage is significantly lower than that of the alternative methods. The method is flexible, in that it is easy to add or change any of the classes. When adding a new class however, care should be taken to ensure that classifications that were correct are not affected in any way. The datasets used in this project could be used for such purpose.

As is the case for all classification methods, it is recommended to calibrate this new algorithm for each country, since the vehicles that are used may differ from those in regular use in The Netherlands. A periodic re-calibration is also recommended to accommodate for new vehicle types, and changed configurations.

## References

**Doorn, R.A. van**, "Voertuigcategorieën", Doc. nr. W-DWW-2000-066, Augustus 2000.

**Kendall M., Stuart A. and J. Ord**, 1983, *The Advanced Theory of Statistics*, Volume 3, IV-th edition, pp 370-375, Charles Griffin & Co Ltd, London.

**Oommen, B.J. and Loke, R.**, "Pattern Recognition of Strings With Substitutions, Insertions, Deletions and Generalized Transpositions", *Pattern Recognition*, Vol. 30, 1997, pp. 789-800.