STUDY OF SAME-LANE AND INTER-LANE GVW CORRELATION

B. Enright, *Department of Civil & Structural Engineering, Dublin Institute of Technology, Dublin, Ireland*
E.J. OBrien & C.C. Caprani, *School of Architecture, Landscape and Civil Engineering, University College Dublin, Ireland*

ABSTRACT

Extensive work has been done over the last two decades on the simulation of traffic loading on bridges. The methodology used is to generate a number of years of simulated traffic and to use extreme value statistics to predict more accurately the characteristic loading for a given bridge. The parameters and probability distributions used in the Monte Carlo simulation must be based on observed sample traffic data. Many previous studies have assumed that there is no significant correlation between the Gross Vehicle Weights (GVW) of trucks in the same lane, or between trucks in adjacent, same-direction lanes.
For this paper, an extensive database of Dutch Weigh-in-Motion data is analysed. Data is collected from two same-direction lanes and is time-stamped to the nearest 0.01 seconds. The statistical characteristics of this set of data are presented, and various techniques are used to establish the nature and extent of GVW correlation.

1. INTRODUCTION

1.1 BACKGROUND

It is well established that traffic loading on many road bridges is considerably less than for the network at large or for roads of the class in which the bridge is located. This can be very useful when bridges fail a capacity assessment by a small margin, as it may cause the bridge to be retained where it otherwise would have needed to be repaired or replaced. Therefore, the load assessment of existing highway bridges is an area where great savings in maintenance budgets are possible.

For 2-lane bridges with traffic travelling in opposing directions, the traffic streams in each direction can be assumed to be statistically independent. Where there are same-direction lanes on the other hand, vehicles may be coming from the same source and their weights may be correlated. For example, there is anecdotal evidence of the existence of overweight convoys such as a crane and a truck carrying its kentledge. Conversely, on such bridges, it is reasonable to expect that only lighter trucks occur in the overtaking lane, due to better mechanical performance.

For this paper, an extensive database of Dutch Weigh-in-Motion data is analysed

1.2 SOURCE DATA

The Dienst Weg- en Waterbouwkunde (DWW) office of the Dutch Ministry of Transport maintain Weigh-in-Motion (WIM) sensors on 3 lanes in one direction of the 6-lane motorway near Woerden in central Holland. Data for truck traffic in the two inner lanes for the 20 weeks period from 7th February to 25th June 2005 were made available to the Bridge and Transport Infrastructure Group in the School of Architecture, Landscape & Civil Engineering in University College Dublin. No data was supplied for the outer lane which trucks are not normally permitted to use.

The data were supplied in a series of files. One set of files contained the following data for a total of 725,897 trucks:

- Vehicle number (unique identifier)
- Date
- Time (to nearest second)
- Speed
- Lane
- Category (type of truck)
- Length

- Individual Axle loads, the sum of which is the Gross Vehicle Weight (GVW).
- Axle spacings

These data files were loaded into a Microsoft® Access database. A second set of log files contained almost 20 million records for many different types of events related to the operation of the WIM sensors. Among these were timestamps to the nearest 0.01 seconds for each truck. Such accurate time-stamps are essential for the modelling of the gaps that occur between same-lane trucks. These timestamps were extracted from the log files and stored with the other truck data by using relational database join operations.

1.3 DATA CLEANING

Data quality issues were identified in consultation with DWW, and the original list of trucks was reduced by eliminating unreliable readings. The criteria used were:

- The timestamp for the truck should be also recorded in the log file so that the more accurate timestamps (to 0.01 s) are available. For various operational reasons, 61,554 trucks had not been recorded in the log files, and were excluded from the analysis.
- The recorded speed should be between 60 and 120 km/h inclusive. Axle weights for trucks travelling at speeds outside this range are not considered to be reliable. This resulted in the exclusion of a further 15,839 trucks.
- The number of axles should be two or more. Some "zero-axle" and "single-axle" trucks were mistakenly registered by the WIM sensors. This resulted in the exclusion of a further 79 trucks.
- The GVW should be 3.5t or greater. 200 trucks in the original list were mistakenly registered by the WIM sensors as having zero GVW, but all of these had already been excluded by applying the first three conditions above.

The number of trucks was thus reduced from 725,897 to 648,425. Of these, 598,292 (92.3%) were in the inner slow lane, and 50,133 (7.7%) were in the overtaking "fast" lane. All subsequent analysis described herein was carried out on this reduced set of clean data.

2. KEY CHARACTERISTICS OF DATA

2.1 GROSS VEHICLE WEIGHT (GVW)

2.1. (a) Overall GVW

Two histograms of GVW distribution in the slow lane are shown below – for 0t to 60t (tonnes) in Figure 1, and for 60t to 170t in Figure 2 using a magnified vertical scale. The first histogram supports the often-used assumption of a bimodal normal distribution, with one peak at 16t and a second peak for fully loaded trucks at 36t. The legal limit for 5-axle trucks in the Netherlands is 50t, with a limit of 11.5t for an individual axle load. It is interesting to note the significant tail of very heavy trucks in the second histogram which supports the view [4] that different models must be used for the general population of trucks and for the tail of very heavy trucks. As would be expected, the tail of heavy trucks in the fast lane (not shown here) is much smaller, with just 89 trucks over 60t, compared with 1,750 in the slow lane, and the heaviest truck observed in the fast lane was 90t, compared with 166t in the slow lane.

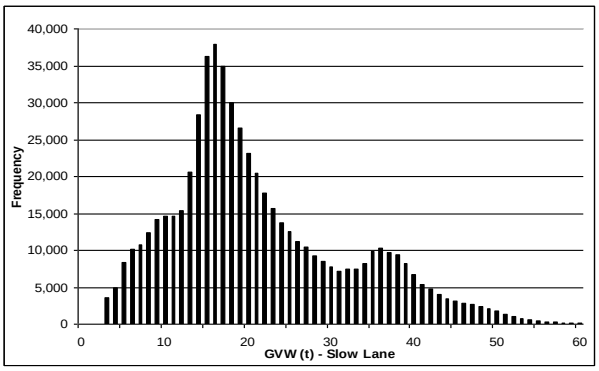

Figure 1. GVW Distribution 0t to 60t – Slow Lane

To illustrate the nature of the very heavy trucks, a summary of all trucks with GVW of 140t or greater is shown in Table 1 (all are in the slow lane).

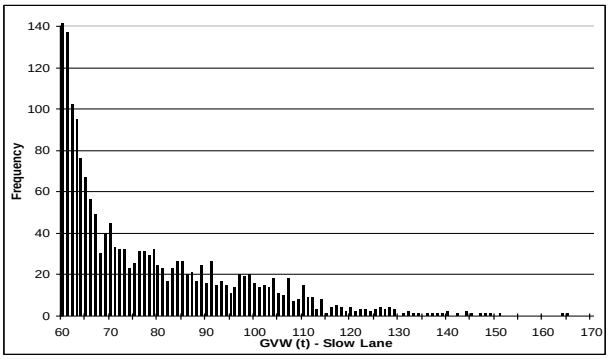

Figure 2. GVW Distribution – 60t to 170t – Slow Lane

Table 1. All trucks over 140t

GVW (t)	Number of Axles	Wheelbase (m)	Speed (km/h)
166	12	28.7	78
165	12	27.3	85
152	13	28.4	80
150	12	28.8	79
148	13	19.5	76
147	12	28.8	81
146	13	36.6	76
145	11	24.8	82
145	13	29.4	80
143	12	28.8	77
140	13	28.3	84
140	13	28.2	86

2.1. (b) GVW by number of axles

Further analysis of the GVW distribution is shown in Figure 3 for 5-axle trucks and in Figure 4 for 9-axle trucks. These illustrate the fact that whereas the distribution of 5-axle trucks is well-behaved, the distribution becomes more fragmented as the number of axles increases. This can be attributed to both sparseness of data and the non-standard nature of trucks with high numbers of axles.

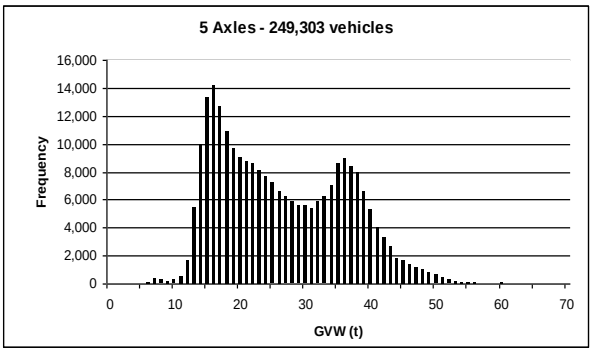

Figure 3. GVW Distribution – 5-axle trucks

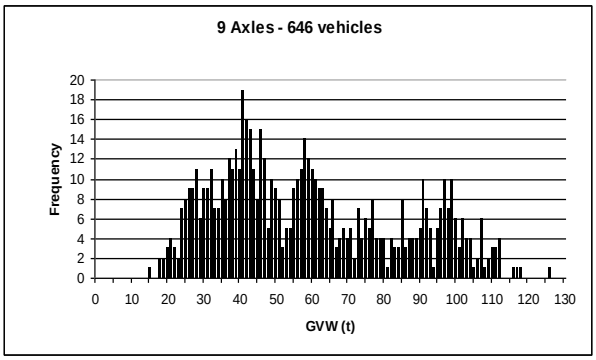

Figure 4. GVW Distribution – 9-axle trucks

2.1. (c) Intra-day GVW variations.

There are significant variations in truck weight over the 24 hours each day in both lanes, as can be seen in Figure 5. There is a sharp peak of 24t in the slow lane in the early morning between 03:00 and 04:00. The daily maximum average hourly flow (not shown here) also occurs around 04:00 – at 353 trucks per hour in the slow lane, and 44 trucks per hour in the fast lane. The average weight dips to 20t by 06:00, and rises back up to nearly 24t by 18:00. In the fast lane, the variation is even more dramatic, from a peak of over 22t at 03:00 to under 16t at 21:00. This intra-day variation in GVW gives rise to positive correlation between the weights of trucks travelling at around the same time of day. This point is discussed in more detail in Section 3.2 on autocorrelation.

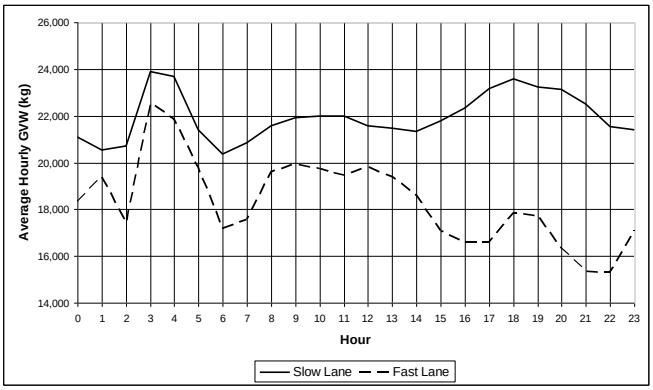

Figure 5. Intra-day GVW

2.2 HEADWAY

Trucks are assigned a timestamp based on the point when the first axle is detected by the WIM sensors. The inter-axle spacings are recorded, and these can be summed to give the wheelbase for the vehicle. The overall length of the body of the truck is also measured by inductive loop detectors. The gaps between successive trucks in the same lane can be measured in different ways. Headway is defined as the time gap in seconds between the first axle of the leading truck and the first axle of the following truck. The headway between vehicles has been used in many studies [1] as the basis for generating simulated traffic arriving on a bridge. The gap may also be measured as the time between the rear axle of the leading truck and the front axle of the following truck. Driver behaviour is influenced by the clear gap between the bodies of the two trucks. However, the measurement of truck body lengths is not particularly reliable, and this lack of reliability is evident in the analysed data.

Figure 6 below show the distribution of headways in the range from zero to 4 seconds. A commonly used assumption [1] is that the coinciding of a number of very heavy trucks in free-flowing traffic represents the critical loading for bridges of relatively short spans (up to perhaps 45m), whereas for longer spans congested traffic is more likely to produce the critical loading. A vehicle travelling at 80 km/h travels 89m in 4 seconds, and in the bridge spans of interest (below 45m) trucks separated by longer headways will not be on a bridge at the same time. Critical multi-truck bridge loading events happen when the headways are very small. Of particular interest in the distribution shown here is the small peak between 0.4 and 0.6 seconds. Previous studies [1] have assumed that the headway distribution drops to zero around 0.7s, whereas these results indicate a small but significant number of "tailgating" trucks. This includes trucks of all weights, and could be significant for bridge loading. The distribution is otherwise very similar to what has been used in recent studies [1] – a negative exponential distribution from 1.5s upwards, with some form of polynomial curve fitting between 0.7s and 1.5s.

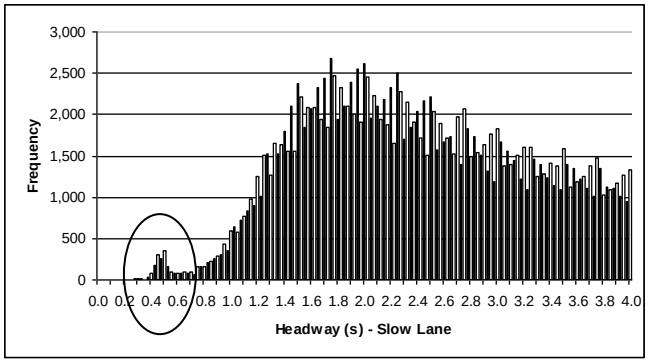

Figure 6. Headway distribution – Slow lane

3. GVW CORRELATION

3.1 CONTOUR PLOTS

The relationship between the leading truck GVW and the following truck GVW in all truck pairs was analysed for trucks travelling together in the same lane, and for pairs of trucks travelling beside each other in both lanes. The statistical model commonly used is the bivariate bimodal normal distribution. The joint probability density function for this theoretical distribution is shown in both 3-D form in Figure 7 and as a contour plot in Figure 8 which show contours of constant probability density. Both of these are based on zero correlation between the two variables.

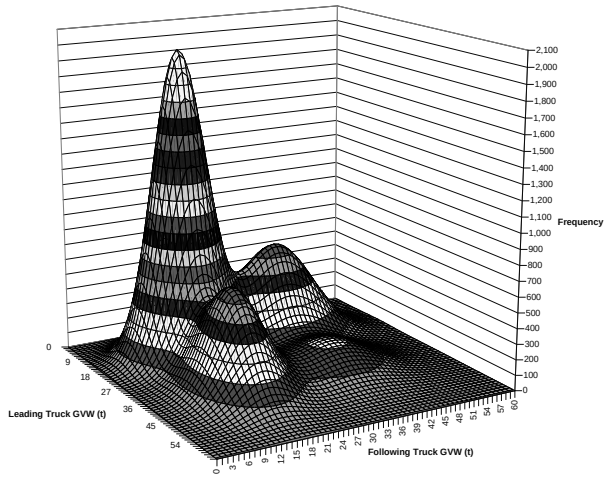

Figure 7. Bivariate bimodal Normal joint probability density

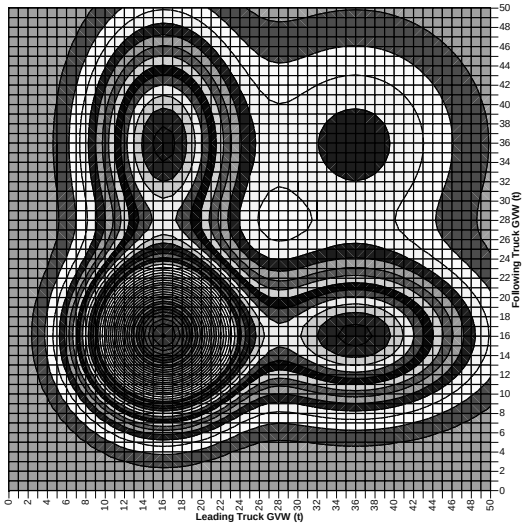

Figure 8. Bivariate bimodal Normal Contour plot

If correlation is introduced into the theoretical data by means of simulation, the shape of the contour plot changes, and this is particularly noticeable for pairs of heavy leading and heavy following trucks where the contours become elliptical rather than circular ("heavy" is defined for the purpose of this study as over 25t). This can be seen in Figure 9 where the data have a 25% coefficient of correlation.

The contour plot for the slow lane at Woerden is shown in Figure 10. This shows that the heavy-heavy zone in the slow lane has a distinctly elliptical shape, which indicates significant correlation between heavy trucks travelling together in the slow lane.

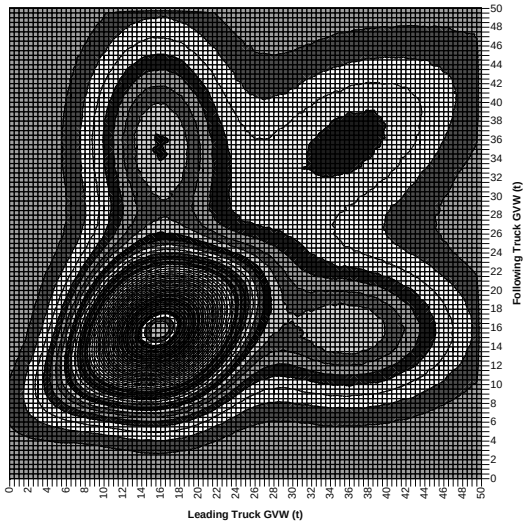

Figure 9. Contour plot – 25% correlation

Similar plots for the fast lane and for inter-lane traffic do not show the same pronounced elliptical shape, and this supports the analysis in the section on autocorrelation that the correlation for heavy trucks in the fast lane and between lanes is not as strong as in the slow lane.

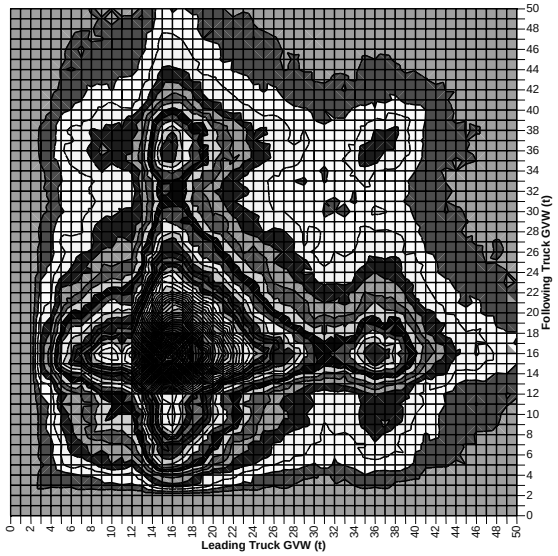

Figure 10. Contour plot – Slow Lane

3.2 AUTOCORRELATION

Autocorrelation is used in the analysis of time series in areas such as economics [6] and signal processing. The term autocorrelation (or serial correlation) denotes the correlation of a random variable with a time-shifted version of itself. A typical time series contains observations of a

random variable X at equally spaced time intervals. The value of the random variable at each time t, X_t, is compared with the value of the variable at time t-s, X_{t-s}, where s is some time lag. The coefficient of correlation is then calculated as a function of the time lag s, and this is referred to as the autocorrelation function:

$$\rho(s) = \frac{E[(X_t - \mu_t)(X_{t-s} - \mu_{t-s})]}{\sigma_t \sigma_{t-s}}$$

A series of truck GVWs can be considered as a time series at randomly spaced time intervals. In this study, the autocorrelation function was calculated using the variable "number of trucks apart" instead of a time lag. The coefficients of correlation were calculated between the weight of each truck (the leading truck) and the truck following it, between the leading truck and the second truck behind it, between the leading truck and the third truck behind it and so on. The results of this are shown for all trucks in the slow lane in Figure 11.

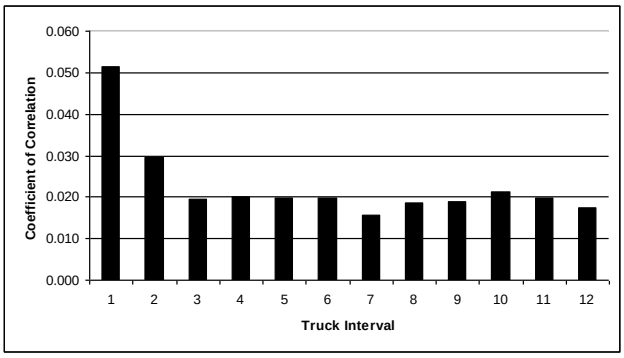

Figure 11. Autocorrelation – Slow Lane

This shows that there is an underlying correlation of 2.0% between trucks travelling at the same time of day, and that there is significantly more correlation (5.1%) between pairs of consecutive trucks. The underlying correlation can be attributed to the intra-day variation in GVW shown earlier in Figure 5, and also to some form of platoon effect whereby heavy trucks tend to be found travelling in groups. Corresponding correlation coefficients for the fast lane are 7.6% (underlying) and 9.7% (intra-pair). Further analysis shows that the correlation in the fast lane is mainly due to lighter trucks. In both lanes, trucks travelling very close together (less than 4s apart) show higher intra-pair correlation (8.7% in the slow lane and 12.4% in the fast lane).

For inter-lane autocorrelation, a slightly different approach was used in calculating the time lag. Each truck in the fast lane was compared first with trucks beside it in the slow lane. This was defined as a truck in the slow lane within 4 seconds in front or behind the one in the fast lane. This time interval was widened to 10, 20, 30, 40, 50, 60, 300, 600 and 1,200 seconds to provide the autocorrelation function. The results are shown in Figure 12. This shows an underlying correlation of 2.1% and an intra-pair (under 4s) correlation of 4.0%. Again, this shows significant additional correlation for pairs of trucks travelling beside each other. This may be attributable to trucks which are travelling together overtaking one another. Average overtaking times for cars has been measured as approximately 8 seconds [5]. Trucks are substantially longer than cars and their relative velocity in overtaking may be lower. A figure of 20 to 30 seconds overtaking time might be considered reasonable for trucks, and this lends support to the suggestion that overtaking may explain the shape of the autocorrelation function for inter-lane traffic.

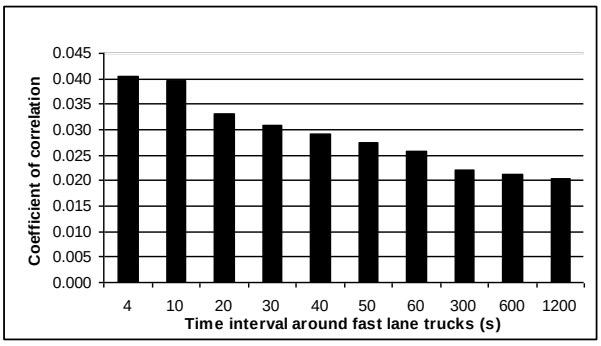

Figure 12. Autocorrelation – Inter-lane

A more detailed analysis was done to establish whether intra-pair correlation is influenced by the absolute weights of both trucks. For different weight thresholds, correlation coefficients were

calculated for pairs of truck where both trucks exceeded the threshold. A 95% confidence interval for the population correlation coefficient (ρ) was calculated using the method described in [2], [3]. The confidence interval depends on both the number of data points (N) and on the calculated estimate for the coefficient (r). A transformed variable z is defined as:

$$z = \frac{1}{2}\log_e \frac{1+r}{1-r}$$

The variable *z* is approximately normally distributed with mean and standard deviation:

$$\mu_z = \frac{1}{2}\log_e \frac{1+\rho}{1-\rho}$$

$$\sigma_z = \frac{1}{\sqrt{N-3}}$$

Using these, a 95% confidence interval for *z* and hence *r* can be calculated. There is a requirement that the two random variables for which the coefficient of correlation is being calculated should be at least approximately possess a joint normal distribution [2], and this is the case here, particularly when correlation is being calculated for pairs of heavy trucks or pairs of light trucks.

The data become sparse as the weight threshold increases, particularly when the much lower traffic volumes in the fast lane are being analysed, and as a result the calculated coefficients become unreliable for higher weight thresholds. The results are shown in Figure 13. The data point plotted here for zero GVW is the coefficient of correlation between pairs of light trucks (where both are under 25t). The 95% confidence interval for the slow lane is also shown. For the fast lane and inter-lane data, the lower bound of the confidence interval drops below zero for weight thresholds above 35t. It is clear that there is a sharply increasing correlation between pairs of trucks in the slow lane as the weights of both trucks increase. This corresponds to the distinctly elliptical shape evident in the contour plot in Figure 10 above. This may well be significant for the prediction of critical bridge loading.

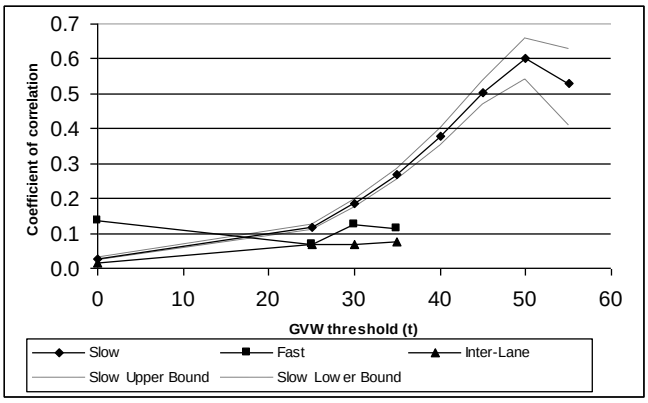

Figure 13. Correlation vs. weight threshold

4. CONCLUSIONS

Some interesting characteristics were identified in the data which will need to be incorporated into future traffic simulation for bridge loading. These include the number of extremely heavy trucks (up to 166t), and the tailgating behaviour of some trucks.

Significant correlation was found between the weights of pairs of trucks. This is particularly true for pairs of very heavy trucks in the slow lane, and further work is needed to quantify the significance of this correlation for bridge loading.

ACKNOWLEDGEMENTS

The authors would like to thank Mr Han van Saan of DWW in the Dutch Ministry of Transport for his assistance in providing and interpreting the WIM data.

REFERENCES

1. Caprani, C.C, 2005 'Probabilistic Analysis of Highway Bridge Traffic Loading', *PhD Thesis, University College Dublin.*

2. Hoel, P.G., 1971 'Introduction to Mathematical Statistics', *Wiley pp 166-169*

3. Fisher, R. A., 1921 'On the probable error of a coefficient of correlation deduced from a small sample' *Metron, 1 (4), 3–32*.

4. Getachew A., OBrien, E.J., 2005, 'Sensitivity of predicted bridge traffic load effects to the tails of truck weight distributions' *5th International*

Conference on Bridge Management, University of Surrey, Proceedings pp 275-282
5. Hegeman, G., Brookhuis, K., Hoogendoorn, S., 2005, 'Opportunities of advanced driver assistance systems towards overtaking' *EJTIR, 5, no. 4 (2005), pp. 281-296*
6. Enders, W. 1995, 'Applied Econometric Time Series', *Wiley*