A cellular model for road traffic

Wojciech Kania
Department of Physics and Astronomy, Manchester University

Harry Lipscomb
Department of Physics and Astronomy, Manchester University

Abstract. —An analysis of traffic flow within single and double lane systems using a discrete automation model is carried out below. Road density dependant optimum speed limits are successfully estimated for these circular road systems of 22mph to 56mph. The improvement in flow when going from the single to double lane models was also inspected and found to be a factor of 1.92 to 1.65 greater, dependant on road speed limit. Finally, a 'funnelling lane' style model was then explored which was seen to have an equilibrium density which changed with speed limit.

1. Introduction

A cellular automata, the first of which was devised by John von Neumann in 1951, is an abstract computational system made from a finite grid of cells. Each of these cells is in a discrete state (both temporally and spatially) and, through a series of simple laws, interacts with its local neighbours to evolve in parallel with other cells at discrete time steps[1]. Given these properties, cellular automation is not only used within the study of traffic flow but in the research of neurons within the brain, ferromagnet production and the arrangements of cellular automata generations can even be used for random number generation[2]. Cellular automation acts as a highly valuable tool for studying complex, chaotic systems of interdependent smaller constituents and is the basis of this papers models.

The layout of this paper is as follows: The theory behind traffic car flow (Sect 2), an outline of the single and double lane models (Sect. 3a and Sect. 3b respectively) followed by a range of results and their relevant analysis. This includes the peak car flow study (Sect. 4a) and the optimization of traffic speed limits for the single lane model (Sect. 4b and Sect. 4c). This is build on through an extension to the double lane model (Sect. 4d) and a funnelling lane variation (Sect. 4e). Finally a conclusion and a look ahead to future ideas is in Section 5.

2. Theory

The primary focus of the study was on the concept of traffic flow (sometimes alternatively named volume), denoted q. This is defined as the total number of vehicles that pass a section of the road per unit time. It is found through the equation

$$q = \frac{\sum_{t=0}^{T} \sum_{t=0}^{L} V_i(t)}{L \times T}$$

where T is the length of time the road was looked at, L is the length of the road and $V_i(t)$ is the velocity of a car in the i$^{\text{th}}$ space on the road (a function of time). This essentially calculates the total movement of road traffic over a length of time and then divides by the length of the road and the total time to find the average vehicles passing a space per time step of the model.

This flow was often plotted against the road density: σ. The road density is simply the total number of cars on the road at any one time divided by the length of the road. It should be expected through previous studies of both traffic and particle flow when plotting q against σ the flow reaches some maximum point where laminar, unimpeded motion of cars is reached. At densities above this point, so called 'phantom' traffic jams may be observed[3] where traffic can greatly slow without any obvious obstacle aside a low random break probability included in the below models. These mean that at very high densities the resulting car flow is low as these jams build up and cars cannot accelerate close to the speed limit imposed on the road.

3. Setup of Models

Both single and double lane models were defined on one-dimensional arrays of length L with open or periodic boundary conditions. Each element of the array represents a short segment of the road of one car length which may be empty or occupied by one vehicle with an integer velocity ranging from zero to v_{max}.

3a. Single Lane Model

The single lane model was defined on a single one-dimensional array. This setup was devised to mimic a circular road which loops around back to the start, conserving total car number.

The simulation of the traffic consisted of consecutive updates of the system, originally devised in *"A cellular automaton model for freeway traffic"* [4]. Each update had four phases which were performed for all vehicles in parallel:

1) Acceleration phase - for all the vehicles with velocity lower than v_{max} and distance to the car ahead larger than $v_{max}+1$, velocity was increased by 1

2) Slowing down phase - if the velocity of a vehicle is larger than the distance ΔL to the next vehicle in front of it, its velocity is reduced to $\Delta L - 1$

3) Random brake phase - the velocity of all moving vehicles may be decreased by one, with probability p, fixed for the duration of the simulation

4) Moving phase - all cars advance v sites. Where an index error occurs and a car leaves the road, a new car is spawned with the same velocity in a location to mimic a circular road setup (this was later modified in Sect. 4e)

Step 3 is essential to the model as it is used to closer mimic a real road. As mentioned in the theory, traffic jams are observed on the roads everyday due to these minor random distractions and breaks of cars. This means that a state of laminar flow with completely unimpeded traffic isn't reached for all road densities making the results more interesting and relevant than if the cars didn't demonstrate this property.

3b. Double Lane Model

The double lane model was defined on two one-dimensional arrays, each treated as an independent single lane system. They interacted with each other through an additional lane changing phase which was added before the acceleration phase. At this stage, a vehicle at position i was allowed to change lane if its velocity was larger than the distance ΔL (to the next car), provided that the road between $i-1$ to $i+(\Delta L+1)$ on the other lane was empty. The improvement in the capacity of the double lane system to demonstrate a larger flow in comparison to the single lane system is analysed in Section 4d below.

There was a range of criteria that could have been used to determine when a car is able to switch lanes. Potentially, the velocity of cars on the opposing lane should also have been considered as to reduce the number of sharp breaks made due to the lane switching mechanism. But, for the road densities used within this study, such a step meant that it was rare cars were able to change lane so this was left out and the more basic requirements above were used. This is a possible future extension that could be analysed.

4. Results and Analysis

4a. Peak Car Flow

First, the total car movement per time interval against road density (σ) was analysed (car movement rather than flow so the length of the road as a variable could also be inspected). The below figures show this for the single and double lane models respectively.

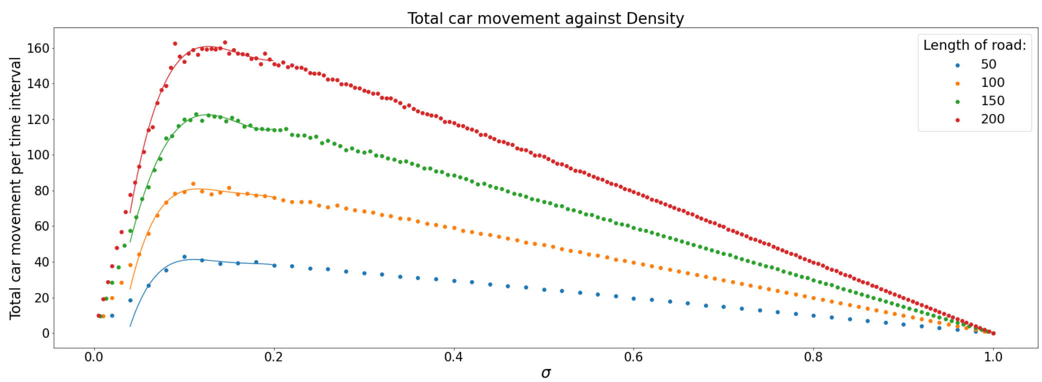


Figure 1. Total car movement against density - single lane model. A random breaking probability of 0.1 and a maximum speed of 10 cars lengths per time interval were used above.

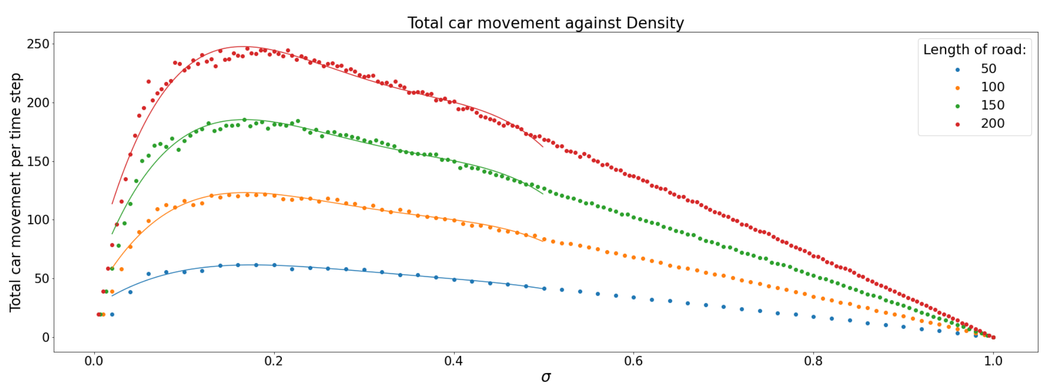


Figure 2. Total car movement against density - double lane model. Identical parameters to Figure 1 were used here.

Figures 1 and 2 above demonstrate how the car flow (equal to the car movement per time step per length of road) peaks at some road density (referenced below as the optimum road density) and then decreases for roads more crowded than this. At too lower densities, all the cars quickly reach maximum speed but this results in a low flow as there are so few cars. To contrast, at high densities there are far more cars braking and queues for the cars to move through, preventing them from forming a high traffic flow. This is in agreement with the theory mentioned in Section 2 above.

To determine the optimum road density for various speed limits, high degree quadratic curves were fit to the data so an accurate maximum could be found. To form the most accurate value, both models were repeated 4 times as this produced an uncertainty on the optimum road density values. For example, for the single lane road with a speed limit of 10 car lengths per time, the density which formed the peak car flow was: 0.116 ± 0.003. The low uncertainty on this value supports the use and effectiveness of the curve fitting method used as the code is reproducing very similar results each run. As observed in Figures 1 and 2, the optimum road density was found to be independent of the length of the road. This makes intuitive sense as the highest average speed of each car should only depend on how close the next car is and not on the total length of the road. The concepts explored above were extended to a range of speed limits as the maximum speed limit imposed on the traffic affects this optimum road density. The relationship between speed limit and the optimum road density is highlighted below.

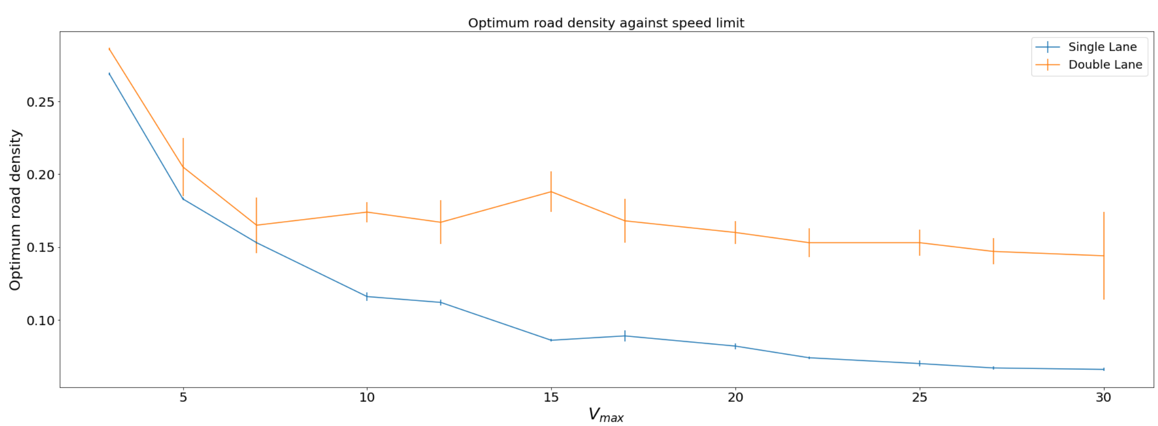
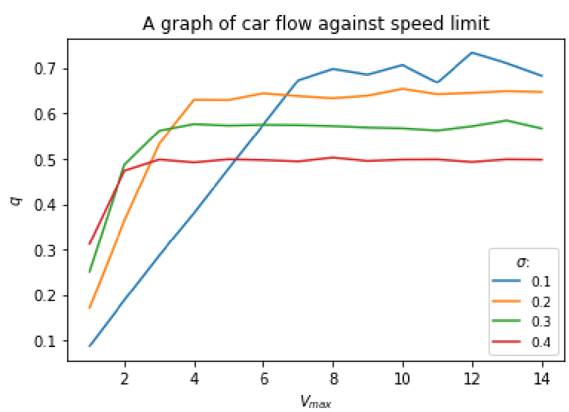


Figure 3. Optimum road density against speed limit. A random breaking probability of 0.1 and a road length of 300 was used.

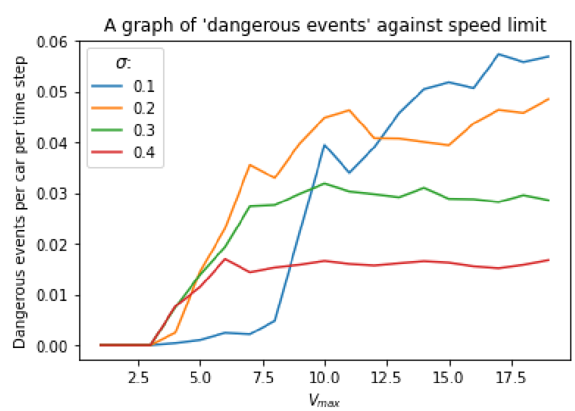
What can be noticed from Figure 3 is that a double lane road copes much more effectively with larger road densities as it can impose higher speed limits for high densities. This is to be expected because in the double lane model slower cars cause less congestion due to the lane changing system allowing faster vehicles to overtake and so cars are able to travel at higher velocities largely unimpeded. However, at densities of over 0.2 both lanes of the double lane model experience frequent traffic jams and so a very slow speed limit must be imposed for both models at such road densities to allow laminar traffic flow.

4b. Optimum Speed Limits and Safety

The optimum speed limit imposed on a road should not only result in a road that flows quickly but safely. Figures 4a and 4b could be used to determine a speed limit which does exactly this. During this modelling, a 'dangerous event' was recorded every time a car was forced to slow down by 3 velocity units per time unit.



(a) Car flow against speed limit and road density.



(b) Dangerous events against speed limit.

Figure 4. Two figures which may be combined to estimate an optimum speed limit for a given road density.

Through Figure 4a it can be seen that car flow increases with speed limit up to a certain point and then remains essentially constant (this is also observed in the 3 vertical contours of Figure 5). At speed limits larger than this, the limiting factor in the overall car flow is not the speed limit but the space for the cars to safely accelerate up to this speed limit. This graph can then be compared to Figure 4b which essentially represents the risks posed to drivers by various speed limits.

In Figure 4b it is seen that dangerous events increase with imposed speed limit. A low car density road is observed to be the safest road at low speed limits but becomes the most dangerous at high speed limits. Here the cars are most likely to reach very high speeds in comparison to high density roads at fast speed limits where cars may struggle to reach fast speeds due to road congestion.

For a body enforcing speed limits for roads with a large volume of traffic ($\sigma \approx 0.4$), a speed limit greater than 3-4 car lengths per time unit seems adequate to achieve maximum car flow whilst also minimising the number of dangerous events and consequently the number of collisions between cars. If the road in question has a low density ($\sigma \approx 0.1$), then a speed limit of roughly 7 car units per time interval provides the perfect balance of car flow and safety for drivers.

4c. A Quantitative Estimate of Speed Limits

To produce estimates of optimum speed limits and ensure the model is applicable to real situations, both the length and time units of the model must be calibrated to reflect real roads. The average length taken up by one vehicle in a jam was approximated as 5 metres hence each distance unit within our model is 5 metres. The units of time within our model were found from comparing the peak car flow values in Figure 5 to measured car flow values within other models or observed maximum traffic flows on real roads.

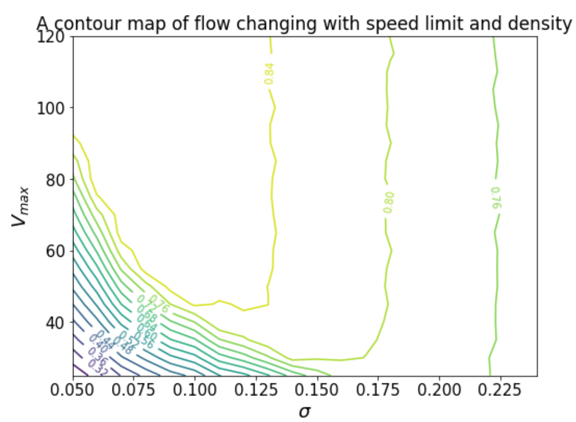


Figure 5. A contour plot to show how density and maximum speed limit affects the flow of cars, note here V_{max} is in units of metres per time step.

By averaging the values from two computer models[5,6] for the maximum road flow for single lane traffic, defined as $f_{observed}$, this gave a value of 2,140 cars/hour. On the other hand, when using real road data from studies[7,8] this value was found to be 1,680 cars/hour. By estimating our model to have a flow of 0.84 cars/time unit, defined as f_{model}, the time units of this model (in seconds), denoted t can be found through

$$t = \frac{3600 \times f_{model}}{f_{observed}}$$

. This gave an interval of time in the model as either 1.41s or 1.80s respectively. The longer time could be used to consider speed limits appropriate for cars with very low reaction time such as autonomous cars (like within the referenced models above) where as the longer time for human drivers. Propagating these through the previously recommended speed limits, a road density of 0.1 cars per space would be best served by a speed limit of roughly 56mph to 44mph. On the other hand, a busy road of density 0.3 to 0.4 cars per space would be best with a speed limit of 28 mph to 22 mph.

4d. Double Lane Variation

The main purpose of the double lane model was to investigate whether adding a second lane to the system would double the maximum car flow. Figure6 below shows the ratio of double to single lane traffic flow for different velocities.

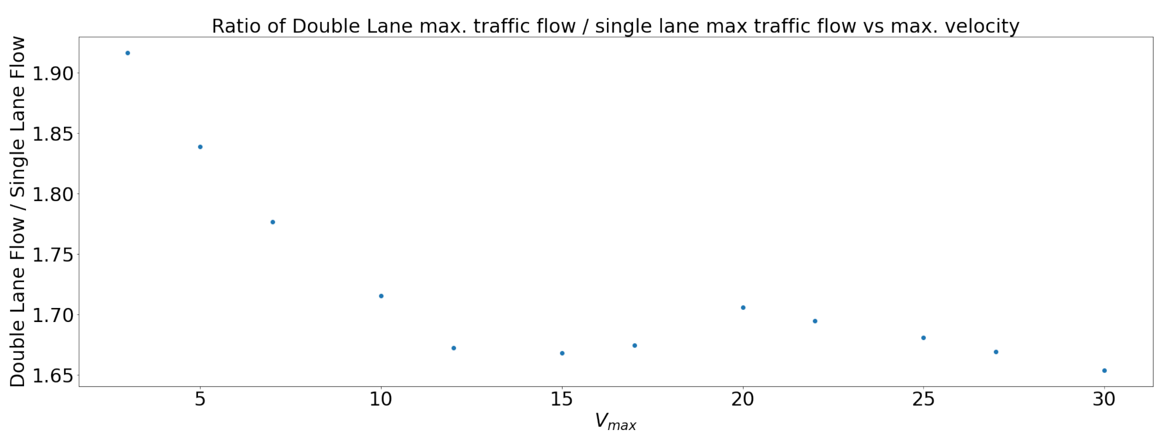


Figure 6. Ratio of double lane max flow/single lane maximum flow vs velocity

Figure 6 shows that for low maximum velocities the total car flow is almost doubled however as the speed limit increases the ratio decreases until it plateaus around 1.67 ± 0.02. The sharp decrease in the effectiveness of the double lane road is caused by the increase of heavy breaks when cars change lanes to avoid slowing down. To investigate this, the frequency of dangerous events was plotted against the maximum velocity in Figure 7.

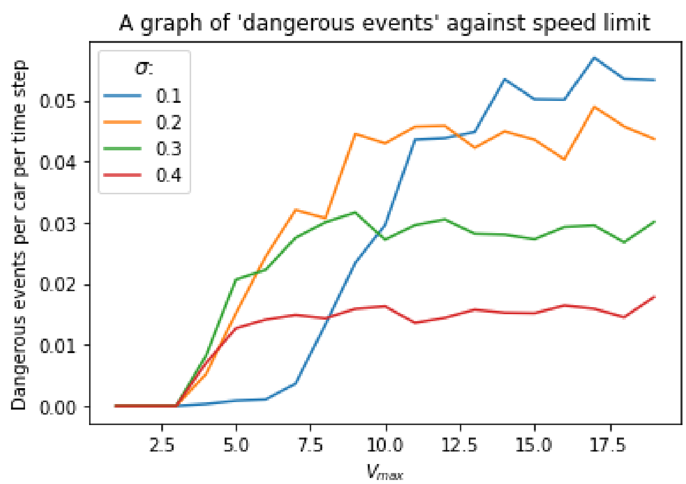


Figure 7. Plot of dangerous events against speed limit for double lane model.

As seen in Figure 7, the average number of dangerous events per car per time step starts increasing when the speed limit exceeds 3 spaces per time step. Dangerous event frequency is much higher for lower road densities because cars more often reach high velocities making them more likely to run into a car which just entered their lane with a lower velocity. For higher densities, this doesn't happen as often as not many cars can reach speeds high enough to require heavy braking should a slow car move in-front of them. Compared to the analogous results for the single lane model, shown in Figure 4, the probability of dangerous events is 2 - 3 times more likely for the double lane road.

4e. Funnelling lane comparison to Circular Road Model

In the previous sections a model where cars were travelling in a loop was analysed. However, when considering a section of road in the real world, cars leave and new ones enter it. The system simulated in this section is a single lane road with cars entering it whether there is enough space at the beginning, which expands to a double lane at the end. To simulate this, the boundary conditions of the model were changed so that cars at the last two spaces disappear and new cars are spawned on the first space every time this first space is empty. This may reflect a large multi-lane road where cars must funnel into a single lane, essentially creating a continuous stream of traffic which may join the road.

The main difference between the two models was in the behaviour of the car density. During the loop model it was held constant as the car number was conserved however for the funneling lane variation it reached equilibrium after a certain length of time as shown in Figure 8.

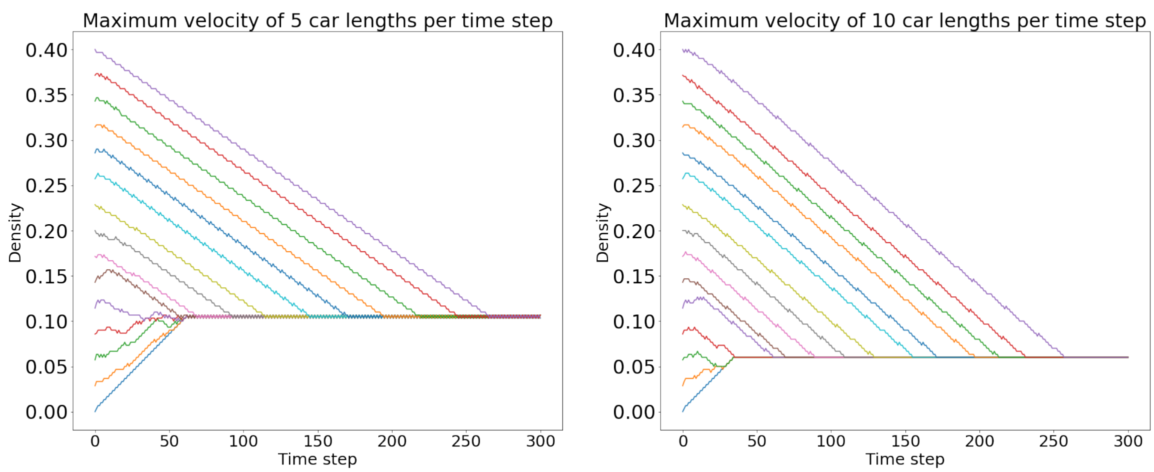


Figure 8. Density of the car spawning model.

Interestingly, the starting density and the length of the road only affects the time it takes for the system to reach equilibrium. This time increases linearly with both length of road and the difference between initial and equilibrium density. The final density for the open road depends only on the speed limit, as shown in Figure 9.

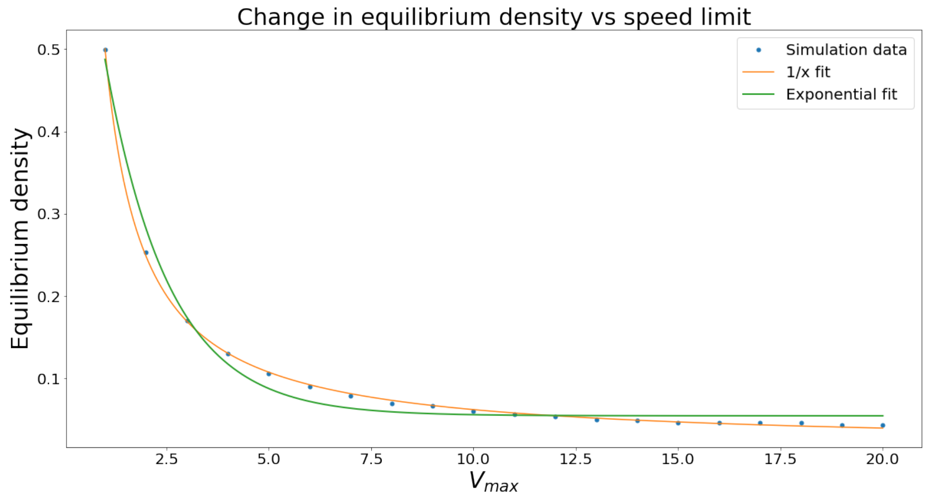


Figure 9. Equilibrium density against speed limit for the funnelling lane model.

The speed limit dependence was expected since the total time a car spends on the road is directly proportional to the maximum velocity. Considering a simple system, an empty road with length L and speed limit V_{max}, each car spends L/V_{max} time steps on the road, which is equal to the total number of cars if they entered the system with velocity V_{max}. However, cars enter the system with 0 velocity (as there's no space in-front) and in the first time step they can only move by one space, hence blocking the following car waiting to join the road. The consequence of this is a decrease in the number of cars on the road by a factor of 2. The equilibrium density, ρ, is thus found by putting these expressions together to give

$$\rho = \frac{L/V_{max}}{2L} = \frac{1}{2V_{max}}.$$

To investigate this dependence, a scatter plot (Figure 9) was created and two different functions: 1/x and e^{-x} were fitted to the simulation data. As predicted by the above equation, the 1/x was almost a perfect fit, however the exponential function was also a good representation of the data. The uncertainty on the equilibrium density was of the order 1/L, so with the length of the road as 400 spaces it was approximately 40 times smaller than the lowest density, therefore negligible.

5. Conclusion and Future Ideas

To conclude, the circular road model of traffic flow was used to demonstrate laminar and turbulent traffic flow. This idea was extended to determine safe speed limits through defining a 'dangerous event' which were found to be between 56mph to 44mph for low density roads and 22mph to 28mph for high density roads. The double lane and funnelling lane models were also explored above. There is a range of potential extensions that could be made to the study. A closer look at the definition of a 'dangerous event' and a stronger link to real world data could be included. Also, an emphasis on how the model could be altered to reflect a human driver (and their reaction time) in comparison to a fully automated car would also be an interesting concept.

6. References

[1] Berto, F. and Tagliabue, J. "Cellular Automata", *The Stanford Encyclopedia of Philosophy*, Spring 2021 Edition

[2] Sarkar, P. "A Brief History of Cellular Automata", *ACM Computing Surveys (CSUR)*, **32**(1)

[3] Flynn, R.M., Kasimov, R.A., Nave, J.C., Rosales, R.R., Seibold, B., Traffic Modeling - "Phantom Traffic Jams and Traveling Jamitons", https://math.mit.edu/traffic/.

[4] Nagel, K., Schreckenberg, M., "A cellular automaton model for freeway traffic", *Journal De Physique I*, 1992, pp2221-2229.

[5] Dervisoglu, G., Gomes, G., Kwon, J., Horowitz, R., Varaiya, P., "Automatic Calibration of the Fundamental Diagram and Empirical Observations on Capacity", 2009.

[6] Muhammad, T., Kashmiri, F.A., Naeem, H., Qi, X., Chia-Chun, H., Lu, H., "Simulation Study of Autonomous Vehicles' Effect on Traffic Flow Characteristics including Autonomous Buses" *Journal of Advanced Transportation*, **2020**, 2020.

[7] Greenberg, H., "An Analysis of Traffic Flow", *Operations Research*, 1959, **7**(1), pp79-85.

[8] Asgharzadeh, M.A., Kondyli, A., "Comparison of Highway Capacity Estimation Methods", Conference: Transportation Research Board, 2018.