Investigating the Causes and Effects of Traffic Congestion

Aman Bhalla[1]★
[1]University of Southampton, University Rd, Southampton SO17 1BJ

15 April 2024

ABSTRACT

Traffic congestion is a complex process that poses economical and logistical challenges to the design and legislation of transportation networks in every nation. In this paper I will aim to explore the dynamics of congestion using traffic simulations. I will employ a cellular automaton model, which uses kinetic theory to model individual cars. This model views individual cars analogously to discrete gas particles that interact with each other. As a result of the high volumes of cars, I will utilise Monte Carlo simulations to predict useful parameters from this simulation. This paper will analyse the effects of congestion through the lens of journey times. I will explore the interplay of random slow down events and speed limits on reducing journey times. I will then understand the underlying causes of congestion by analysing the effects of driver ability on traffic flow. The results will provide insights for both highway engineers and drivers to ensure the efficient flow of traffic in complex transportation systems.

1 INTRODUCTION

Understanding and optimising traffic systems are a vital component to upholding the economic and social security of a nation. Transportation infrastructure is responsible for transporting the population, goods and military equipment. An important tool to improving these systems can be traffic flow simulations and models. For example, traffic flow models have allowed for infrastructure improvements in critical natural disaster situations. In these events, a reliable and efficient traffic system is necessary to ensure the safe transport of the population during the event (Tyagi et al., 2009). Predictive traffic flow models in Lebanon found that earthquakes can result in network connectivity losses of 70%. Their model was able to reveal defects in the road network, allowing first aid services to better visualise and prepare for these weaknesses (Sarkissian et al., 2020).

Another application of traffic flow models is in urban development. Congestion in motorway and local systems proves to be a serious challenge to vehicle mobility. Congestion is ultimately a result of traffic demand exceeding the capacity of a motorway. This can occur if there is a lack of roads/lanes, however, there are many factors including speed limits and random events which can affect this. Congestion has resulted in increased emissions, fuel consumption and a steep increase in travel time (Tyagi et al., 2009). In the UK, 70% of the workforce commutes to work by car, with the average driver spending 124 hours per year stuck in congestion. This has resulted in an estimated cumulative cost due to congestion of £307 billion (between 2013-2030). Subsequently, it is imperative for cities to employ innovative solutions to combat these effects. As such, technological innovations like traffic flow models and simulations can provide a better image of which solutions prove most effective in tackling congestion (INRIX 2014).

By applying physical theories to traffic systems, we can generate models to further understand the underlying causes of congestion. There are two primary methods in which physical scientists do this: hydrodynamical models and kinetic theory (Nagel and Schreckenberg 1992). Hydrodynamical models aim to derive equations for the evolution of macroscopic quantities of traffic flow analogous to mass, momentum and energy in conventional hydrodynamics. As such, they do not focus on the individual cars but on traffic density instead. Changes in densities of the flow are often approximated by shock-waves or some nonlinear diffusion equation (Bonzani 1999). Hydrodynamical models can be applied to understanding the effectiveness of traffic control measures. For example, using an unsteady-state flow, researchers found that at low traffic densities, speed bumps and traffic lights effectively controlled traffic without disturbing the free flow (Smirnov et al., 2014).

Kinetic theory aims to model traffic flow analogously to gas flow. Cars are treated similarly to individual gas particles which interact with each other. A limitation of this model is that it does not capture the continuous adaptive nature of human drivers as the velocity of each car can only change discretely with each time step (Tampère et al., 2003).

In this paper I will use kinetic theory to model the flow of traffic, with rules defined by a cellular automaton model that splits the road into individual cells each occupied by a single car. This method is fully outlined in section 2.3. I will aim to investigate some causes and effects of congestion. To understand the causes, I will look at how driver ability effects traffic flow. Driver ability can vary from slow and reactionary drivers to aggressive drivers who do not obey the speed limit. To understand the effects of congestion, I will investigate how journey times are affected by congestion. I will specifically look at how random slow down events and speed limits affect journey times. This paper will begin by outlining the methods used to achieve and analyse this model, including the cellular automaton model, random numbers and Monte Carlo simulations. I will then outline the results of the simulation, including errors and plots, followed by a discussion on the real-world validity of these predictions.

2 METHOD

The following section will outline some fundamental techniques to create the traffic simulation, including a brief overview of random numbers and Monte Carlo simulations. I will then outline the cell automaton function which regulates the motion of each car, and an explanation of how I carry out the simulation. Finally, I will outline plotting techniques to visualise the various parameters analysed in this project.

2.1 Random Numbers

A fundamental technique in traffic flow models is the use of random number generators. Random number generators produce a set of numbers derived from specific types of number distributions. Some examples include uniform and Gaussian distributions. A uniform distribution produces numbers with equal probabilities of being chosen. For all choices, the distribution can be visualised by a horizontal line of equal probability. Gaussian or normal distributions have a probability distribution about the mean of the data, as such a visualisation of the graph takes on a bell curve shape. Data closer to the mean has higher frequencies compared to data at the extremities (Haahr; Chen 2023; Chen 2024).

In this paper, I repeatedly use random number generators to define the positions of cars on the road and the probability of random slow down events. As I use the python "random" module, random values are produced from a uniform distribution. However, these values are not truly random. Python uses a pseudo-random generator to produce these numbers known as the Mersenne Twister. This method is capable of producing 53-bit floats with a period of 2^{19937} - 1, however due to its deterministic nature, it cannot produce completely random numbers (Python Software Foundation).

2.2 Monte Carlo Simulations

Monte Carlo Simulations are a method of obtaining the likelihood of some uncertain event using repeated random sampling. The general method begins by identifying the independent variables used to make the prediction, and the dependant variable we are trying to predict. The independent variables are then selected from a probability distribution. The simulation is then run repeatedly, attempting to try every combination to predict the dependant variable (IBM).
In this paper, I use Monte Carlo simulations when estimating traffic flow and journey times, by taking repeated, random samples.

2.3 Cellular Automaton Model

To create a traffic flow simulation, a set of rules are required to outline the motion of each car with respect to other cars and the conditions of the road. In this paper I will follow the cellular automaton model developed by Nagel and Schreckenberg in 1992. This method splits the road into cells that can be occupied by a single car at a time, where the motion of each car is governed by 4 rules.

I have implemented this approach in python, however the following method can be applied to any relevant coding language.

2.3.1 Initialisation

To use this approach, I begin by initialising the road with a function taking values of the road length (L), number of cars (N) and the random seed used. I model a straight road divided into L cells using an array of length L. The contents of each cell represents the velocity of a car and the index of each cell represents the position. An unoccupied position is defined by an impossible value of velocity. In this simulation I arbitrarily take this to be -1. An unoccupied position is defined by velocity 0, which can increase to a value of v_{max}. Initially, I assign N amount of stationary cars to random positions in the road. For plotting the simulation, I also initialise 2 arrays that store unique, random RGB values for each car, and integer id values for each car. Id values start from 1 and increase by 1 until the last car on the road. The road is also taken to be circular, this means that a moving car at the end of the road will return to the beginning. I will now define the rules of movement for each car, in each time step.

2.3.2 Rules of Movement

The following rules define how the velocity of each car updates in each time step. As such I iterate over the whole length of the road array for each step, except until rule 4, where the final positions of the cars are assigned due to their velocity.

1. **Acceleration** If a car has a velocity below v_{max}, v will be increased to v + 1.

2. **Slowing Down** For any car with v >= 0, the code will check for the closest car ahead at some site i + d. I use the modulo operator to calculate the remainder index for cars near the end of the road, to account for the circular nature of the road. If v >= d, v is decreased by 1. This ensures that cars do not accelerate when there is a car ahead of them.

3. **Random Events** This rule aims to model any random events, such as irrational driving, that may slow down a driver. If a car is moving (v > 0) its velocity has a chance to decrease by 1, given by a probability p.

4. **Cars Move** Finally, the velocity of each car is added to its current position. Once again, this is done using the modulo operator to calculate the remainder for any cars near the end of the road, where the remainder becomes the index from the beginning of the road. Cars are then assigned to their final position.

This model assumes that all cars have the same size, speed limits remain constant along the road and that cars accelerate linearly. The model does not contain multiple lanes, junctions or different types of vehicles.
The rules of the model are contained in a function which is later simulated over many time steps. The function also assigns the unique RGB colour values and ids of each car to new positions in parallel to rule 4. To simulate the above rules, I iterate over t time steps, calling the cell automaton function in each step. In each iteration I store the current road in a 2d array, which is used later for plotting Time vs Position. The unique colour values are also stored in a 3d array for plotting (each colour value is a 1d array).

2.3.3 Extension: Variation in Driver Ability

An important cause of congestion can be a variation in the speed and reaction of drivers. To implement this idea in the Cell Automaton Model, I introduced 2 inherent variations in driver ability; hesitancy and aggression.

Hesitant drivers have more extreme reactions to slow down events resulting in significant decreases to their velocities. To do this, I assigned a skill value (S) for each driver, this value defines how much each car slows down when faced with a random event. This changes rule 3 of the cell automaton model, so any car with velocity greater than S - 1 that encounters a random event, will be slowed down by S. To ensure v does not decrease below 0, S < v_{max}. For example, with v_{max} = 5, if a driver is travelling at v = 4

with S = 4 and encounters a random event, it will slow down all the way to a stop (v = 0) in one time step.

Aggressive drivers have a faster acceleration than other cars and are not affected by slow down events. To do this, I assigned each car with a value of aggression (A), A = 0 being a regular driver and A = 1 being aggressive. This impacts both Rule 1 and Rule 3 of the cell automaton model. If A = 1, drivers with $v < v_{max}$ will have $v_{max} + 5$, as such aggressive drivers are able to go over the speed limit. For A = 1, drivers will also not be affected by the random slow down events defined by p.

2.4 Parameters of Interest

2.4.1 Journey Time

To produce values of journey times, I used a function which calculates the time for all cars to complete one loop of a defined length. I began by creating an array that took on 0 values if all cars had not looped or 1 if they had looped. The index of this array correlates with the integer car id values. To calculate the journey time, I checked whether each car had enough velocity to return back to the beginning of the road, and if true, its respective value in the looped array became 1. To get a final value for journey time (T), I kept adding each time step until every car had looped (all 1 values in the looped array), and then taking an average over all the cars. This means that the distance of the journey is equal to the length of the road.

2.4.2 Traffic Flow per Time Step

A useful result of traffic simulations and an important consideration in understanding the effects of congestion, is traffic flow. This defines the amount of cars passing a specific point in each time step. In this code I arbitrarily choose the end of the road, although any point along the road can be treated as equivalent. While simulating each second, I also define a counter which checks how many cars in the road have enough velocity to surpass the end. If they do, the flow counter is increased by 1, and an average value over all the time steps is returned. The flow count can be impacted by the maximum velocity, road density and event probability. These influences will be further explored in the results section.

2.5 Plotting

2.5.1 Time vs Position Graph

A useful way of visualising the traffic flow, would be to plot a Time vs Position graph. To do this I convert the 2d array defining all the roads over all time positions, into 1 array which assigns each car with a value representing its time position. The unoccupied -1 values are now converted into 0s to plot each line of the road. To create a colour map for each car, I carry out the same procedure with the 3d colour array.

2.5.2 Journey Time Graph

To understand how average journey times (T) change with different parameters, I plotted T vs v_{max} and T vs p. I employed a Monte Carlo Simulation to estimate these values over many trials. To do this I defined an array for either v_{max} or p. I iterated over this array, running the simulation for a certain amount of trials for each value in the array, whilst returning the relevant value of journey times. In each iteration, I took the average of journey times over all the trials, and the standard deviation of these average values. I then fit this data using a non-linear least squares method. This allowed me to extract parameters and their errors. To get a better understanding of the suitability of the fit, I used an R^2 method, which determines the amount of variance between the independent and dependant variables of the model. The error bars use the standard deviation of the average journey times produced over all the trials.

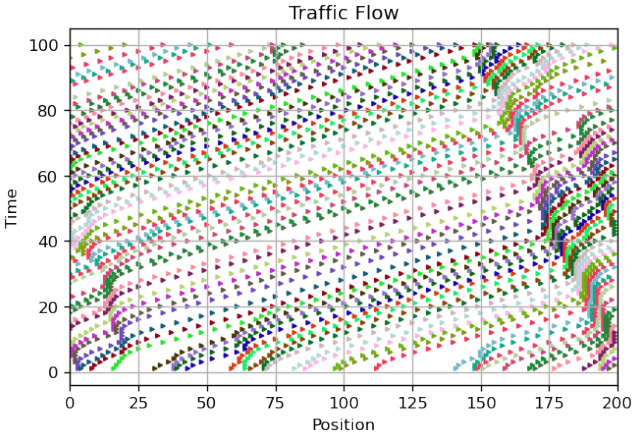

Figure 1. Time vs Position graph visualising a traffic simulation with road length (L) = 200, v_{max} = 5, event probability (p) = 0.25 and a number of cars (N) = 30

2.5.3 Traffic Flow versus Traffic Density

To understand more about the evolution of traffic flow with certain parameters, I plotted traffic flow versus traffic density. This follows a similar method to plotting journey time, except the independent variable I iterate over is the density of cars (N/L) whilst plotting the traffic flow. The methods of fitting and error bars are identical.

3 RESULTS

3.1 Qualitative Analysis

I will begin the results section by taking a qualitative view on the evolution of some of the fundamental characteristics of the model.

From Figure 1, we can see that for an event probability of 25 (percent), traffic jams begin to form. In this case they are positioned near the beginnings and ends of the total length. The positioning of these jams is a result of the random assignment of positions and the random probabilities of slowing events. In Figure 1, I used a random seed of 41 chosen arbitrarily. The jams occur when any car begins to randomly slow down, forcing the cars behind to slow their speed as a result of rule 3 in the cell automaton model. Cars not affected by the jam follow a linear progression, moving at v_{max} = 5 whilst maintaining an equal separation with cars ahead and behind until they loop round to the beginning of the road. For p = 0 cars would simply accelerate up to v_{max} and then maintain a linear progression.

An animation for the traffic flow in figure 1 can be found here.

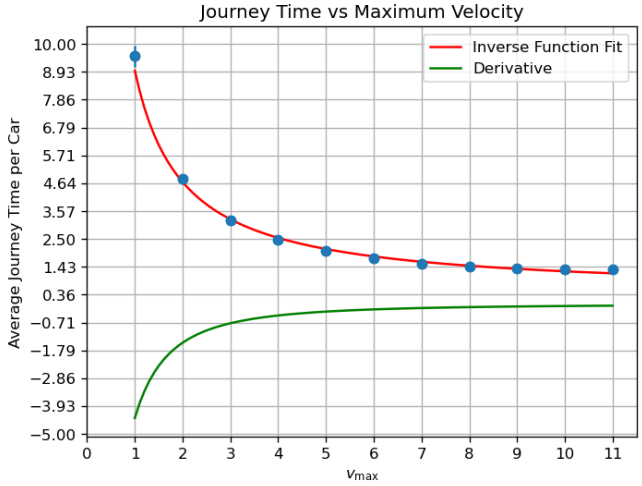

Figure 2. Evolution of Journey Times with v_{max}. L = 200, t = 400, N = 20, p = 0. Simulated for 100 trials.

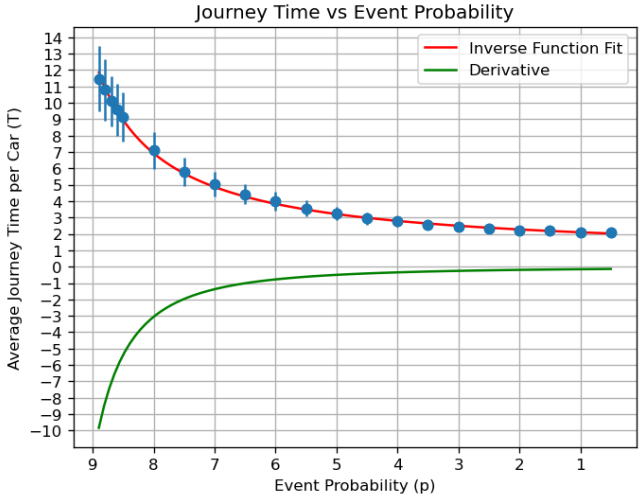

Figure 3. Evolution of Journey Times with p. L = 200, t = 400, N = 20, v_{max} = 5. Simulated for 100 trials.

3.2 Investigating Journey Times

Congestion can result in significant decreases to the average journey time of a car. In this section I will investigate whether v_{max} or p has a greater effect on the average journey time per car (T) All values will be quoted to 3 significant figures unless otherwise specified.

Figure 2 and Figure 3 show the evolution of Journey Times with both maximum velocity and event probability. I did this for velocities ranging from 2 to 11 and event probabilities ranging from 5% to 90%. Attempting to go above 90% led to a significant increase in journey times which meant most cars were unable to complete a loop of the road for the parameters defined in the figures. These ranges were also chosen to capture the most variation in the graph (when it is not at an asymptote). To maintain the scale of the plots between v_{max} and p, I divided all p values (after producing them) by 10.

I selected the total time of simulation to be = 400, to allow cars at high event probabilities to complete a full loop of the road. The choice of 100 trials was made to optimise errors, with an average standard deviation of 0.13 (2 s.f) for maximum velocity and 0.72 (2s.f) for event probabilities. Note, the errors for T vs p are significantly skewed at higher values of p.
Looking at T vs v_{max} I find that as v_{max} increases, average journey times tend towards a constant around 1.37s (2s.f) and as v_{max} decreases, T tends towards infinity. As such, I have used an inverse function to fit the graph defined by the equation:

$$T = \frac{a}{v_{max} + b} \tag{1}$$

In this context, the physical interpretation of parameter a is a scaling factor that represents how much variation the independent variable causes in T. Parameter b represents the horizontal shift in these values and can indicate when the independent variable begins to have an effect on the dependant variable. a = 8.581 ± 0.316, b = 0.399 ± 0.0588 and R^2 = 0.994, indicating a reliable fit. I also took the derivative of the function to understand the rate of change of each

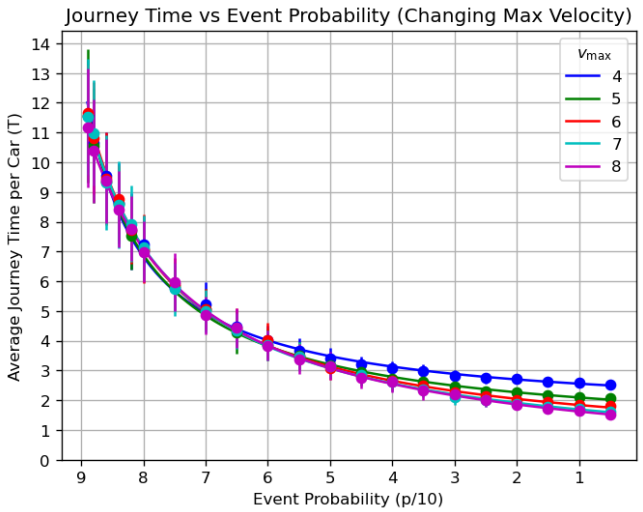

Figure 4. Evolution of Average Journey Times with event probability for different values of v_{max}. L = 200, t = 400, N = 20. Simulated for 100 trials

case, this took on the form of:

$$\frac{dT}{dv_{max}} = \frac{-a}{(v_{max} + b)^2} \tag{2}$$

To make the comparison between both cases easier, I have inverted the x axis for T vs p to show how journey times change as event probability decreases from 90%. Looking at the trend of Journey Time vs event probability, we can see that journey time tends to a constant (≈ 3.4s) as p decreases. However, as event probability increases, the average journey time tends toward infinity. This means T vs p also has an inverse relationship similar to T vs v_{max}. In this case the parameters are a = 12.6 ± 0.642, b = 10.0 ± 0.07015, and R^2 = 0.998. To get a more accurate fit, I introduced a parameter c which vertically shifts the fit closer to its asymptote, c = 0.698 ± 0.0688.

Looking at the derivative of both graphs, we can see that increasing v_{max} past 6 or decreasing p below 40% results in the graph to plateau. This results in no significant impact in the average journey time per car.
Comparing the growth parameter (a) between both cases, we find that event probability seems to have a bigger effect on average journey times, where $a_{event} \approx 1.56 a_{max}$. However, it is important to consider the fact that there may be some relation between the effects of p and v_{max}, which could result in a differing impact on the average journey time. As I have only looked at the case for $v_{max} = 5$, the growth parameter of p may change at different v_{max}.

To further understand this, In Figure 4 I have plotted T against p for 5 different values of v_{max}. By calculating the growth parameters for each of these fits, I find that increasing the maximum velocity causes the reduction of event probability to have an even bigger effect on T. The growth factor increases by approximately 200% between $v_{max} = 4$ and $v_{max} = 8$. At $p > 75\%$, changing v_{max} seems to have no impact on the growth. Therefore, when there are a lot of slow down events, increasing the maximum velocity will not be able to further reduce journey times.

3.3 Effects of Driver Ability on Traffic Flow

In this section I will be investigating how the varying driver abilities defined in section 2.3.3 affect traffic flow. Looking at the Time vs Position graph in Figure 5, we can get a qualitative idea of these effects. Red cars indicate aggressive drivers, who are able to drive over the speed limit, and accelerate as much as they can in each time step. They also do not react to slow down events. This can be seen in the linear progression of aggressive driver groups. They remain tightly packed with each other whilst maintaining a constant speed of $v_{max} + 5$. Green drivers obey the regular rules of the cell automaton model. Yellow cars indicate hesitant drivers who react extremely to slow down events. As we can see, aggressive and regular drivers seem to follow these slow drivers closely as queues form behind them. An animation for the traffic flow in figure 5 can be found here. I will now outline the effects this has on traffic flow (amount of cars passing the end of the road per time step).

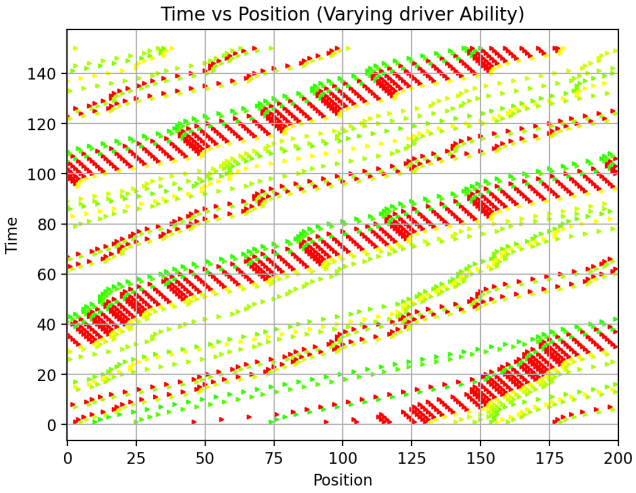

Figure 5. Time vs Position graph visualising a traffic simulation with varying driver ability. Red = Aggressive (above the speed limit, no slow down events). Yellow = Hesitant (extreme responses to slow down events). Green = Normal (regular response to slow down events). L = 200, $v_{max} = 5$, $p = 0.25$ and N = 30.

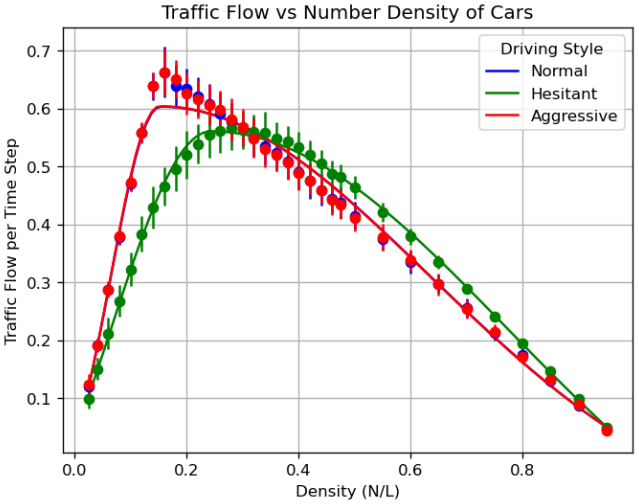

Figure 6. Traffic Flow (number of cars passing the end of the road per time step) vs Number Density (number of cars/Road length) for 3 driving styles, Aggressive (above the speed limit, no slow down events), Hesitant (extreme responses to slow down events), Normal (regular response to slow down events). $v_{max} = 5$, L = 200, t = 200, $p = 0.20$, 140 trials.

3.3.1 Traffic Flow vs Number Density

To understand how driver ability affects traffic flow, I plot Traffic Flow (F) vs Number Density (d) for cars obeying a) Regular cell automaton rules, b) Hesitant drivers and c) Aggressive drivers. The results are displayed in Figure 6. The simulation ran for 140 trials and t = 200 to reduce errors, $v_{max} = 5$ and $p = 0.20$. A key result from these plots is the peak in traffic flow at lower densities, occurring at $d_{normal} = 0.156$, $d_{hesitant} = 0.249$ and $d_{aggressive} = 0.156$. I used a skewed normal distribution to fit each function, however, the fits do not accurately capture the peaks so the values quoted are from the data. An interesting result, is that aggressive driving seems to have no effect on the position or amplitude of the peaks, where $F_{aggressive} = 0.663 \pm 0.0409$, $F_{normal} = 0.662 \pm 0.0439$. This has the implication that driving above the speed limit does not increase or decrease traffic flow. However, hesitant driving resulted in a decrease in traffic flow, where $F_{hesitant} = 0.566 \pm 0.0371$. This shows a percentage decrease of $\approx 15\%$ in traffic flow for hesitant drivers occurring at number densities $\approx 60\%$ higher than regular and aggressive driving. This was simulated at $p = 0.20$ as large values increased the errors significantly, however larger values of p would simply decrease traffic flow for all abilities and move the peaks to higher densities.

4 DISCUSSION

In this section, I will outline interpretations and applications of the results, a comparison to real world data, and some limitations of the model.

To allow for comparisons with real world data, I will assume parameter values to be roughly $v_{max} = 5 = 120km/hr$, one cell = 7.5 metres and one time step = 1 second. These scalings were approximated in (Nagel and Schreckenberg 1992) and seem to follow a rough agreement with the real world.

A primary effect of congestion is its impacts on journey times. In section 3.2, I investigated how v_{max} and p changed journey times, and which parameter had a bigger effect. I find that increasing v_{max} decreased journey times but only up until a constant $\approx 1.37s$. This is likely a result of the saturation of the road. Although cars are able to travel faster, there are still a finite amount of unoccupied cells on the road. This means the road eventually does not have enough capacity for cars to make use of the full speed limit. Cars would also not be able to travel outside of the circular road to get past this constraint. Looking at the derivative of this relationship, I found that increasing v_{max} past 6 does not result in a significant reduction in journey times. Using the real-world approximations, this is equal to 144 km/hr. A study on cars in Queensland, Australia showed that increasing speeds from 60km/h to 65 km/h saved 1 minute on journey times. However, at higher limits of 100km/h, going to 110km/h only reduced journey times by 2.2 to 5.5 minutes (BigRoad 2020). This data generally correlates with the result I found, that higher speed limits do not result in a linear decrease in journey times.

I also found that decreasing the chances of random slow down events below 40% seemed to have less of an impact on decreasing journey times. As this is a significant proportion of drivers, there may be implications that driving below the speed limit can be encouraged for the safety of drivers, as it does not have significant effects on decreasing journey times. For example, recommendations from RAC (Royal Automobile Club) state that speed limits are a maximum and that to ensure safety, cars should drive at lower speeds in the right conditions (RAC 2020).

Another important result was that attempting to reduce random slow down events had an almost 1.5x effect on the growth parameter compared to increasing the speed limit. This is further evidenced by the fact that at larger event probabilities ($p > 75\%$) increasing the speed limit has no significant effect on reducing journey times. In the real world, random slow down events can be a result of braking, lane changing, weather or simply poor driving. As such, from my results it is crucial to mitigate these factors in order to optimise journey times. For example, policies from the highway code state that in congested systems, drivers should not unnecessarily change lanes. This is because unnecessary lane changes can result in unsafe conditions where other drivers are forced to slow down (GOV.UK 2015).

In the second section of results, I investigated some causes of congestion, specifically how driver ability affected traffic flow. Looking at the distributions, traffic flow seemed to peak at lower densities. Although increasing the number of cars may increase traffic flow, it causes a significant amount of slow down events, as such the relationship is non-linear and traffic flow peaks at a low/medium density. These results can be evidenced by a meta-analysis which found that the relationship between traffic volume and accidents follow a non-linear relationship (Hesjevoll et al., 2016).

Upon simulating this relation for hesitant drivers, I found a $\approx 15\%$ decrease in traffic flow, occurring at number densities $\approx 60\%$ higher. This is a result of the extreme reactions of hesitant drivers to slow down events. This reduces the amount of cars able to complete a loop of the road, as such a higher amount of cars is required to produce a peak in traffic flow.

When I simulated a combination of hesitant and aggressive drivers, I noticed tailgating. This occurred when slow drivers would hold up queues of faster cars behind them. Upon plotting traffic flow vs number density of cars, I found that flow rates were not impacted by aggressive driving and stayed similar to drivers who obeyed the cellular automaton model. This is because aggressive drivers who resorted to tailgating were simply held up at the rate of regular drivers going at the speed limit.

However, this conclusion is not supported by real-world data due to the limitations of the model. Ideally, the model should incorporate multiple lanes, as faster drivers would be able to overtake slower drivers. This would result in an increase in traffic flow, rather than staying constant as seen in this model. The model could also be improved by including junctions and roads with fixed amounts of cars entering. Adding non-linear acceleration would also capture the reality of actual driving.

5 CONCLUSION

In this paper, I completed a comprehensive analysis in understanding the causes and effects of congestion on traffic flow. By employing kinetic theory, I was able to understand the interplay between speed limits, random slow down events and road density in large traffic systems. Looking at the effects of congestion on journey times, I was able to understand the importance of speed limits on journey times. In specific, increasing speed limits to higher values do not provide significant returns on reducing journey times, which is backed up by real-world data. On the contrary, reducing random slow down events such as lane changes and random breaking were more important in reducing journey times. When analysing the causes of congestion, I found that lower/medium traffic densities were able to maximise traffic flow, in order to balance slow down events. A conclusion unique to this model was that aggressive drivers going over the speed limit did not increase traffic flow. However, in real-world scenarios, lane changes would account for this possibility. Ultimately, the findings have provided some useful insights for highway engineers and drivers to ensure the reduction of congestion and efficiency of traffic flow.

6 ADDENDUM

The following addendum outlines applications of monte carlo simulations, random numbers and random walks. An outline of random numbers and Monte Carlo simulations have already been provided in the report, however I will provide an explanation of random walks in the addendum. There will also be 3 applications in academia, and 3 outside academia for each technique.

6.1 Monte Carlo Simulations

6.1.1 Monte Carlo Simulations in Chemistry, Statistical Physics and Biochemistry

In chemistry, Monte Carlo simulations have been used to simulate molecular modelling. This is because the method is able to generate a variety of configurations of complex macro-molecular systems. The configurations are acquired by applying random perturbations to each system, ensuring they are large and energetically feasible. Each configuration occurs under specific thermodynamics conditions. Although the Monte Carlo simulations cannot explain the time evolution of the system, they can provide information on the probabilities of thermodynamic observables, such as free energy (Paquet and Viktor

2015).

In physics, large-particle systems with many interactions require numerical methods to derive analytical solutions to each system. As a result of the increased complexity in computation, Monte Carlo simulations have become widespread in many sub-fields, from quantum mechanics to statistical physics. In statistical physics, Markov chains are often used to define the evolution of states that a system may undergo. Markov chains are stochastic processes (a sequence of time-evolving random variables). A core principle of Monte Carlo methods is the use of stochastic to obtain an uncertain variable, making them highly suitable for Markov chains. Additionally, within the phase space of a system, there are a large yet finite amount of states which can be easily calculated with a Monte Carlo approximation, opposed to a conventional approach (Kastner 2010).

Biochemical models often describe complex metabolic systems, which require non-linear regression analysis of data. Samples are often stochastic, as a result Monte Carlo simulations are a suitable method. In understanding enzyme kinetics, (the efficiency of a catalyst in facilitating a chemical reaction), Monte Carlo simulations have been used to estimate specific biochemical parameters within the reaction with a high degree of accuracy (Tenekedjiev et al., 2022).

6.1.2 Monte Carlo Simulations in Finance, Medical Imaging and Supply Chain Networks

In finance, Monte Carlo simulations have been used due to the high complexity of numerical computation. Numerical methods have been used in testing portfolios, risk analysis and sensitivities. For example, in the evaluation of security prices. These can include stock prices, options prices or commodity prices (raw materials). These prices are defined by underlying state variables. Monte Carlo simulations are used to generate random samples for these variables to ultimately model sample paths of these variables which represent the evolution of these prices over time (Boyle et al., 1997).

In medicine, Monte Carlo simulations have been used to simulate particle emission processes that produce medical images with high realism. Conventional deterministic simulators are only able to approximate a synthetic image using additive models. Monte Carlo models are able to simulate real noise, similar to actual data. They have also been used in MRI and CT images to create digital representations of real objects of patients or small animals. These simulations can also incorporate real clinical data to achieve a more realistic result (Bert and Sarrut 2022).

Supply chains are networks occurring in stages, such as supplying raw materials, manufacturing, wholesale, retail and finally fulfilling the need of a customer. Whilst designing such systems, reliability proves to be a crucial variable which is often difficult to optimise. Monte Carlo simulations can be used to estimate the reliability of different supply chain networks, by computing the probability of a functional system given the probabilities of failure. In this application, Monte Carlo simulations prove faster than exact estimation whilst being more efficient (Ozkan and Kilic 2019).

6.2 Random Numbers

Random Numbers are a set of numbers derived from a specific type of distribution, such as a uniform or Gaussian distribution (Haahr).

6.2.1 Random Numbers in Quantum Mechanics, Statistical Mechanics and Astrophysics

Random number generation is an essential technique in many applications. However, when choosing a generator, the sophistication, true randomness and speed of the generator are all important factors to consider. Most software is only capable of producing pseudo random numbers due to the deterministic nature of the physical hardware. To get truly random generators, researchers have leveraged the truly random nature of quantum mechanics. These methods aim to measure quantum fluctuations in vacuums (Jacak et al., 2021).

In Statistical Mechanics, random numbers are used to consider the motion of particles in an ideal gas. The Maxwell-Boltzmann distribution aims to understand the distribution of molecular velocities and speeds. The velocity distribution is defined as the fraction of molecules with velocities in a specific direction, and is related to the Boltzmann factor. Any component of molecular velocity is represented by a Gaussian distribution. This distribution can then be used to estimate quantities such as the average velocity, root mean squared speed, kinetic energy and the maximum particle velocity (Blundell and Blundell 2006).

In Astrophysics, random numbers are used to model energy transport within stars. This is because particles/photons in stellar interiors exhibit random thermal motion. This energy can be transported by a process called radiative diffusion. In this process photon energy is scattered and absorbed in random directions, slowly diffusing to the extremities of the star as they are produced. This process can take millions of years to have an effect on the luminosity of stellar bodies (Pols 2011).

6.2.2 Random Numbers in Finance, Cryptography and Politics

In finance, random number generators are used in the efficient market hypothesis, an underlying feature of financial theory. This outlines that asset prices are a result of all relevant information, implying that outperforming the market is not possible. Random number generators can be used to create models that predict how fast new information will have an effect on asset pricing. This analysis can also explain the efficiency of this process (Moews 2023).

In Computer Science, random numbers have been used in cryptography. This is the process of developing algorithms which encrypt information as it is transmitted. This means the data can only be accessed by someone with the relevant permissions to decrypt the data. To decrypt an algorithm, a key is required. These keys can be produced using random number generators to create a strong key. This can be done using a true random number generator or a pseudo-random number generator, where the generator must be able to create random numbers at speeds capable of producing keys simultaneously (Smirnoff and Scholten 2019).

In politics, random numbers have been used to promote participation in political events. For example, in jury duty, participants are selected at random to serve as a juror. This has the benefit of ensuring fairness in, eliminating any bias from government officials. It can also promote diversity to eliminate bias in selecting specific groups. This also aims to foster more individual participation from the general public, and diffuse significant power dynamics (Carson and Martin 1999).

6.3 Random Walks

Random Walks are a method of understanding random motion in a variety of fields, from physical sciences to finance. A principle

example of random walks is Brownian motion, originally observed in the random motion of pollen grains in a fluid, a result of the collision of fast-moving atoms with liquid molecules (UCLA).

6.3.1 Random Walks in Biology, Quantum Mechanics and Brownian Motion

In Biology, random walk processes can be used in modelling the movement of organisms, cells and animals. For example, to model the paths of animals, a correlated random walk model can be used. This means there is a correlation between steps in the walk, resulting in a tendency to a specific direction over the full time evolution of the random walk. This can be directly applied to the movement of animals, which often tends to a singular direction (Codling et al., 2008).
Random Walks have also been used in Physics and Computer Science using quantum random walks. Computer Science often neglects physical characteristics of devices. Ultimately, understanding the physical laws behind any device allows for a better understanding of the behaviours of that device. Quantum computation aims to bridge this gap. Some specific applications of random walks in this field include, quantum circuits and algorithms. In conventional random walks, states evolve stochastically. In quantum walks, the random nature of each step is a result of superposition of states or a collapse of the wave-function (Kempe 2003; Venegas-Andraca 2012).
In Diffusion, Brownian Motion is a principle example of random walks, originally observed in the random motion of pollen grains in a fluid. It was found to be a result of collisions of fast moving atoms with liquid molecules. In the movement of water, the self diffusion coefficient has been used to describe the distance a water molecule can move in a time step. In this case, the distances are pulled from a Gaussian distribution. As a result of Brownian Motion, diffusion has an equal probability of occurring in all directions, meaning it is isotropic (UCLA; Roberts and Verma 2013).

6.3.2 Random Walks in Finance, Computer Science and Pharmacy

In Finance, random walk theory has been used to identify patterns in financial variables, such as stock price and the foreign exchange rate. Due to the spontaneous nature of share price fluctuations, they can also be evaluated using Brownian motion. Another example, comes from Britain's economic equity markets where analysis of periodic changes in spot cotton and wheat prices found discrepancies in the average prices were uncorrelated. This is supported by the random walk theory, where the price of one financial product is unrelated to another (Chen 2022).
In Computer science, random walks have been used to construct peer-to-peer networks. These are networks where each device can take on the role of a client or server. Random walks significantly improve over conventional methods to construct such networks. Constructing P2P networks requires the independent, random sampling of nodes in the network. By using stochastic processes and taking samples from steps in random walks, you can achieve results similar to independent sampling (Gkantsidis et al., 2004).
In Pharmacy, random walks have been used in drug re-purposing. This is the process of finding alternate applications for therapeutic drugs. Using data sets that associate specific drugs with genes and diseases, it is possible to create recommendation systems. This mechanism is fuelled by Markov chains and is significantly more efficient at complex computational tasks compared to state of the art approaches (Castiglione et al., 2023).

REFERENCES

Big Road (2020). Speeding Doesn't Save Time, Study Shows - BigRoad. [online] bigroad.com. Available at: https://bigroad.com/blog/speeding-doesnt-save-time-study-shows/[Accessed 15 Apr. 2024].

Blundell and Blundell (2006). Concepts in Thermal Physics WW. www.academia.edu. [online] Available at: https://www.academia.edu/40054642/Concepts_in_Thermal_Physics_S_Blundell_K_Blundell_Oxford_2006_WW [Accessed 15 Apr. 2024].

Bonzani, I. (2000). Hydrodynamic models of traffic flow: Drivers' behaviour and nonlinear diffusion. Mathematical and Computer Modelling, 31(6-7), pp.1–8. doi:https://doi.org/10.1016/s0895-7177(00)00042-x

Carson, L. and Martin, B. (1999). Random selection in politics. [online] Available at: https://www.bmartin.cc/pubs/99rsip.pdf. Castiglione, F., Nardini, C., Onofri, E., Pedicini, M. and Tieri, P. (2023). Explainable Drug Repurposing Approach From Biased Random Walks. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 20(2), pp.1009–1019. doi:https://doi.org/10.1109/tcbb.2022.3191392.

Chen, J. (2024). Uniform Distribution Definition. [online] Investopedia. Available at: https://www.investopedia.com/terms/u/uniform-distribution.asp.

Chen, J. (2023). Normal Distribution. [online] Investopedia. Available at: https://www.investopedia.com/terms/n/normaldistribution.asp.

Chen, Y. (2022). Random walk: statistical data analysis and application in finance. 2nd International Conference on Applied Mathematics, Modelling, and Intelligent Computing (CAMMIC 2022). doi:https://doi.org/10.1117/12.2639464.

Codling, E.A., Plank, M.J. and Benhamou, S. (2008). Random walk models in biology. Journal of The Royal Society Interface, 5(25), pp.813–834. doi:https://doi.org/10.1098/rsif.2008.0014.

Der Sarkissian, R., Abdallah, C., Zaninetti, J.-M. and Najem, S. (2020). Modelling intra-dependencies to assess road network resilience to natural hazards. Natural Hazards. doi:https://doi.org/10.1007/s11069-020-03962-5.

Gkantsidis, C., Mihail, M. and Saberi, A. (2006). Random walks in peer-to-peer networks: Algorithms and evaluation. Performance Evaluation, 63(3), pp.241–263. doi:https://doi.org/10.1016/j.peva.2005.01.002.

GOV.UK (2015). The Highway Code - General rules, techniques and advice for all drivers and riders (103 to 158) - Guidance - GOV.UK. [online] www.gov.uk. Available at: https://www.gov.uk/guidance/the-highway-code/general-rules-techniques-and-advice-for-all-drivers-and-rider~:text=You%20should%20follow%20the%20signs [Accessed 15 Apr. 2024].

Haahr RANDOM.ORG - Introduction to Randomness and Random Numbers. [online] www.random.org. Available at: https://www.random.org/randomness/#:~:text=When%20discussing%20single%20numbers%2C%20a [Accessed 15 Apr. 2024].

Hesjevoll (2016). Effect of traffic volume on road safety. [online] Available at: https://www.roadsafety-dss.eu/assets/data/pdf/synopses/Effect_of_traffic_volume_on_road_safety_23112016.pdf.

IBM (n.d.). What is cryptography? | IBM. [online] www.ibm.com. Available at: https://www.ibm.com/topics/cryptography.

IBM (2023). What is Monte Carlo Simulation? | IBM. [online] www.ibm.com. Available at: https://www.ibm.com/topics/monte-carlo-simulation.

INRIX (2014). Traffic Congestion to Cost the UK Economy More Than £300 Billion Over the Next 16 Years. [online] Inrix. Available at: https://inrix.com/press-releases/traffic-congestion-to-cost-the-uk-economy-more-than-300-billion-over-the-next-16-years/.

Jacak, M.M., Jóźwiak, P., Niemczuk, J. and Jacak, J.E. (2021). Quantum generators of random numbers. Scientific Reports, 11(1). doi:https://doi.org/10.1038/s41598-021-95388-7.

Kastner, M. (2010). Monte Carlo methods in statistical physics: Mathematical foundations and strategies. Communications in Nonlinear Science and Numerical Simulation, 15(6), pp.1589–1602. doi:https://doi.org/10.1016/j.cnsns.2009.06.011.

Kempe, J. (2003). Quantum random walks: An introductory overview. Contemporary Physics, 44(4), pp.307–327. doi:https://doi.org/10.1080/0010715031000110776.

Moews, B. (2023). On random number generators and practical market efficiency. Journal of the Operational Research Society, pp.1–14. doi:https://doi.org/10.1080/01605682.2023.2219292.

Nagel, K. and Schreckenberg, M. (1992). A cellular automaton model for freeway traffic. Journal de Physique I, 2(12), pp.2221–2229. doi:https://doi.org/10.1051/jp1:1992277.

Pols (2011). Stellar evolution. [online] www.astro.ru.nl. Available at: https://www.astro.ru.nl/~onnop/education/stev_utrecht_notes/ [Accessed 15 Apr. 2024].

Python Software Foundation (n.d.). random — Generate pseudo-random numbers — Python 3.8.2 documentation. [online] docs.python.org. Available at: https://docs.python.org/3/library/random.html.

RAC (2020). [online] Rac.co.uk. Available at: https://www.rac.co.uk/drive/advice/legal/speed-limits/.

Roberts and Verma (2013). Brownian Motion - an overview | ScienceDirect Topics. [online] www.sciencedirect.com. Available at: https://www.sciencedirect.com/topics/pharmacology-toxicology-and-pharmaceutical-science/brownian-motion [Accessed 15 Apr. 2024].

Smirnoff and Scholten (2019). The Use of Random Number Generators in Cryptography. [online] www.cryptomathic.com. Available at: https://www.cryptomathic.com/news-events/blog/the-role-of-random-number-generators-in-cryptography#:~:text=The%20role%20of%20RNGs%20in.

Smirnov, N., Kiselev, A., Smirnov, N., Nikitin, V., Silnikov, M. and Manenkova, A. (2014). Hydrodynamic traffic flow models and its application to studying traffic control effectiveness Hydrodynamic traffic flow models and its application to studying traffic control effectiveness.

Tampère, C.M.J., van Arem, B. and Hoogendoorn, S.P. (2003). Gas-Kinetic Traffic Flow Modeling Including Continuous Driver Behavior Models. Transportation Research Record: Journal of the Transportation Research Board, 1852(1), pp.231–238. doi:https://doi.org/10.3141/1852-28.

Tenekedjiev, K.I., Nikolova, N.D. and Kolev, K. (2011). Applications of Monte Carlo Simulation in Modelling of Biochemical Processes. [online] PubMed. Available at: https://www.ncbi.nlm.nih.gov/books/NBK402828/#arts.mode1_ch4.S1 [Accessed 15 Apr. 2024].

Tyagi, V.V., Swaroop Darbha and Rajagopal, K.R. (2009). A review of the mathematical models for traffic flow. International Journal of Advances in Engineering Sciences and Applied Mathematics, 1(1), pp.53–68. doi:https://doi.org/10.1007/s12572-009-0005-8.

UCLA Random walks. [online] Available at: https://www.math.ucla.edu/ biskup/PDFs/PCMI/PCMI-notes-1.

Venegas-Andraca, S.E. (2012). Quantum walks: a comprehensive review. Quantum Information Processing, 11(5), pp.1015–1106. doi:https://doi.org/10.1007/s11128-012-0432-5.

This paper has been typeset from a TEX/LATEX file prepared by the author.