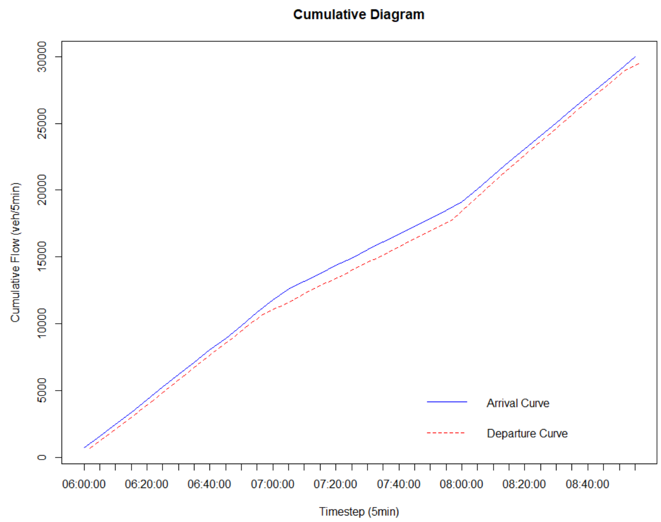
Problem 1

(a) Cumulative diagram:

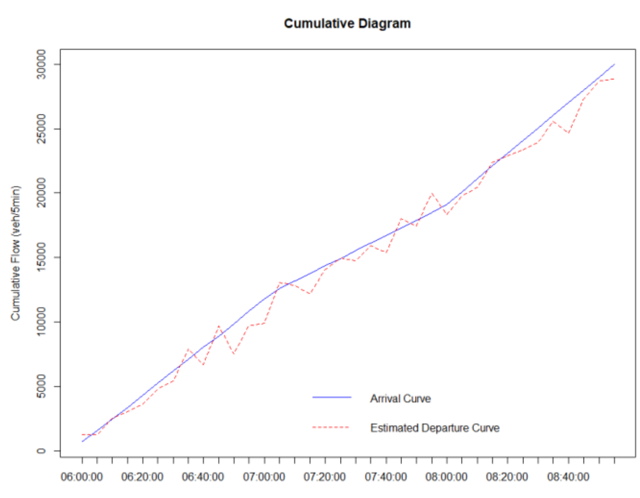


(b) Definition:

The x axis is the time period being observed, whereas the y axis is the cumulative counts of vehicles in this time period. There are two curves in the diagram. One represent the cars entering the road section of interest and the other represent the cars exiting this road section. Empirically, if the conservation law holds in this road section, the rate that cars leave the road section is less than the rate of entering the section. Then in the diagram the arrival curve is always above and has steeper slope than Departure Curve.

(c) the start time period when the cumulative flow of upstream sensor is 900 more than that of the downstream sensor is "2018-05-14 07:06:00".

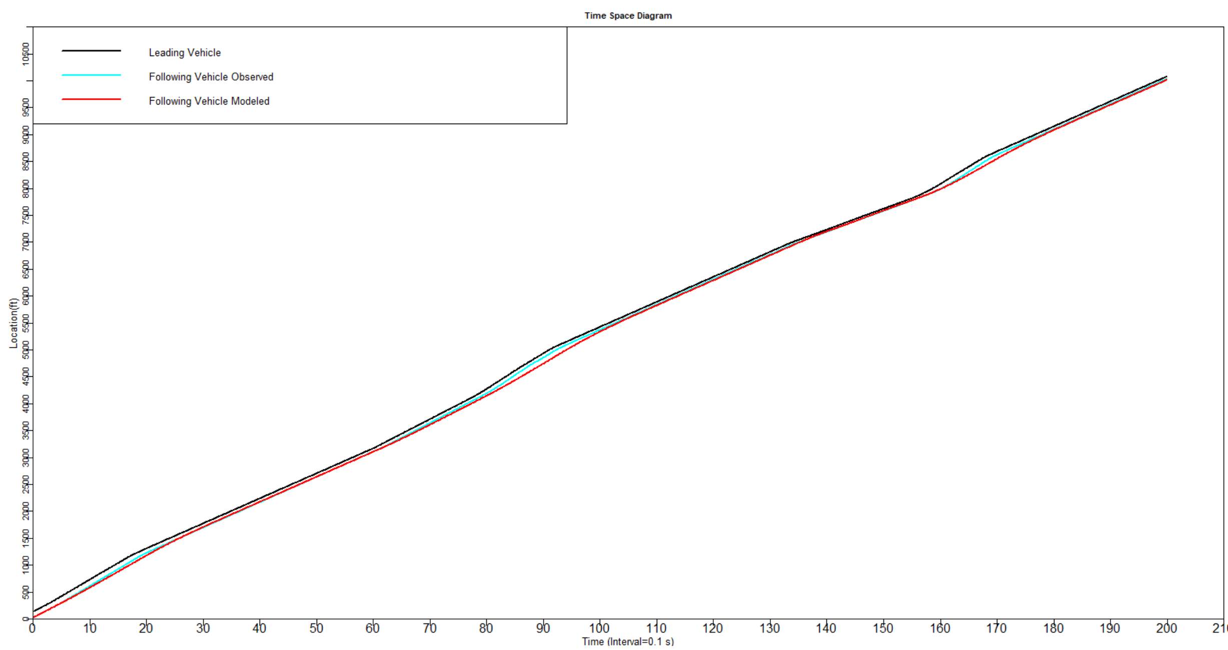
(d) Plotting arrival curve and estimated departure curve:



Problem 2:

(a) Please refer to the R code. The optimal parameters are as follows:
Time interval: 0.1
Reaction time: 1.5
m: 0.1
L: 1
sensitivity coefficient: 10

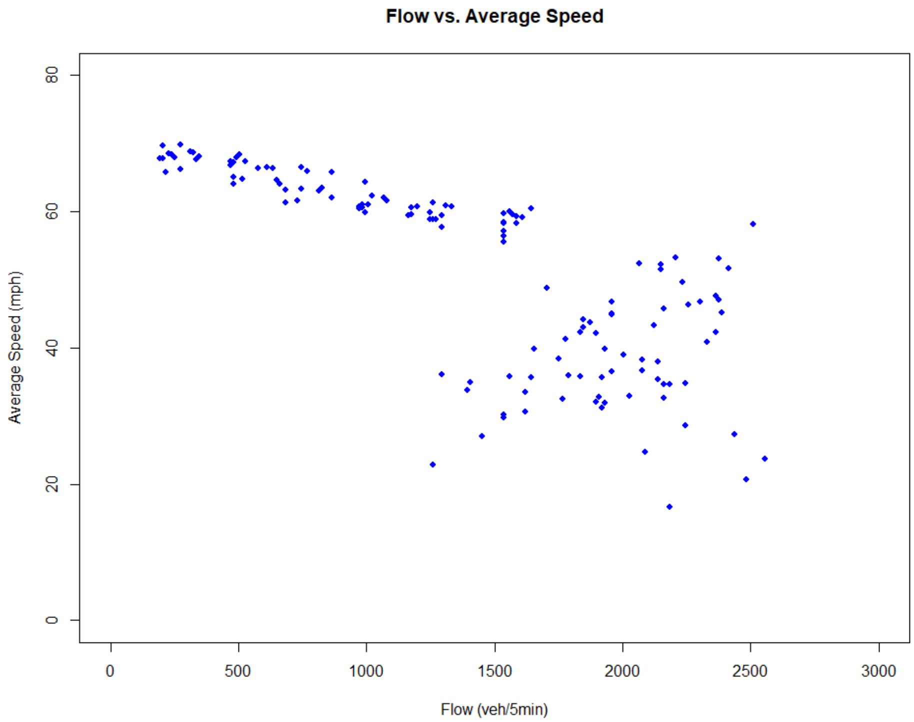
(b) Plotting:



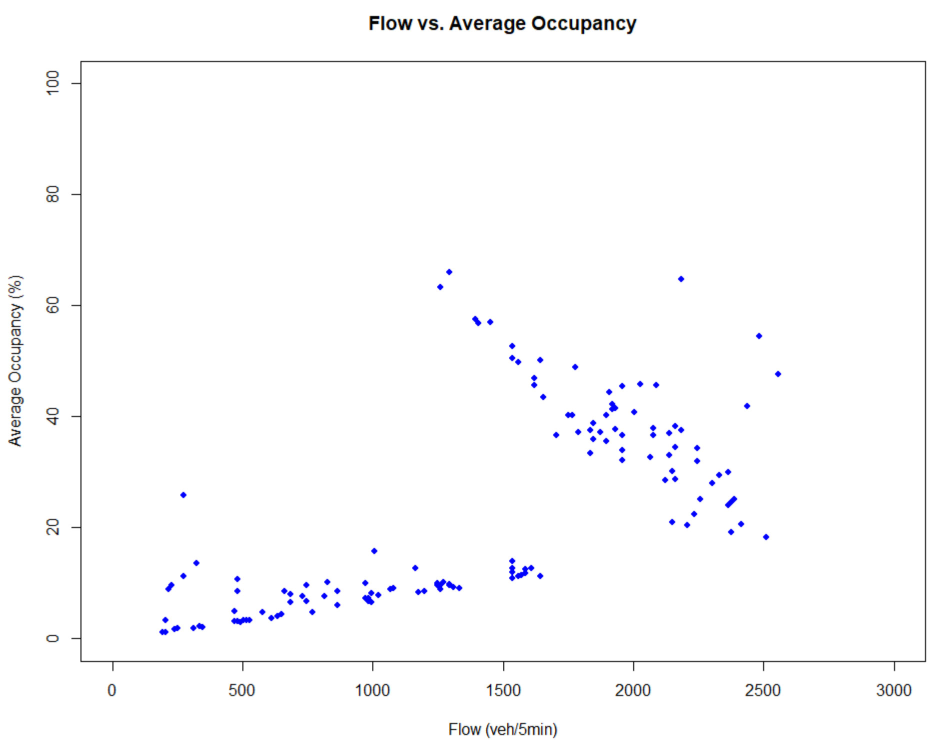
(c) Comparison: Generally speaking the modelled trajectory successfully imitate the trend of Observed one. However, the distance between modelled following car and leading car is larger than the real observation. This is because the modelled trajectory is subject to a much more strict condition of acceleration. In the model the following car only accelerate when the relative speed between following and leading car is larger than 0, whereas in a real world situation, the following car driver may probably accelerate based on his/her own judge which may not be congruent to the modelling situation.

Problem 3:

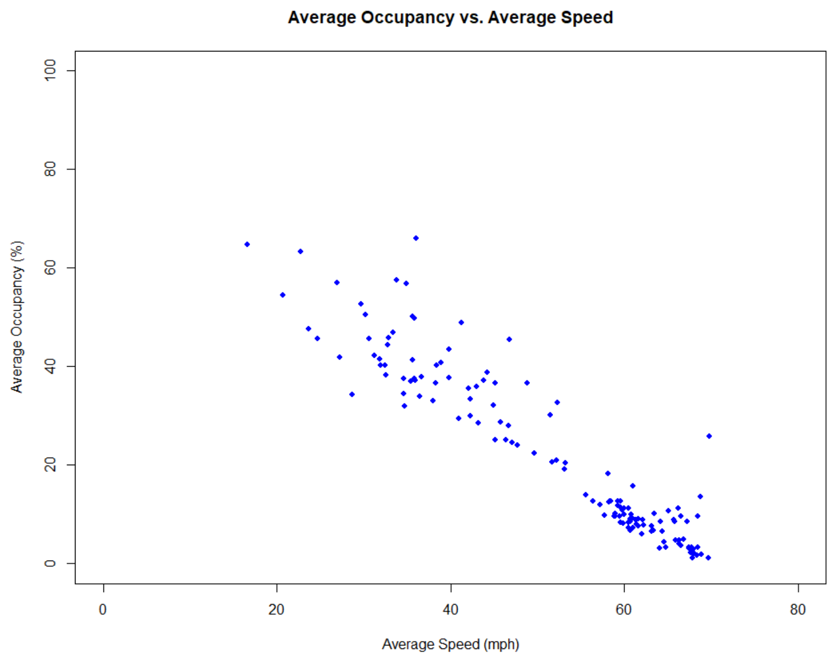
(a) Detector 3:
Flow vs Speed:



Flow vs occupancy:

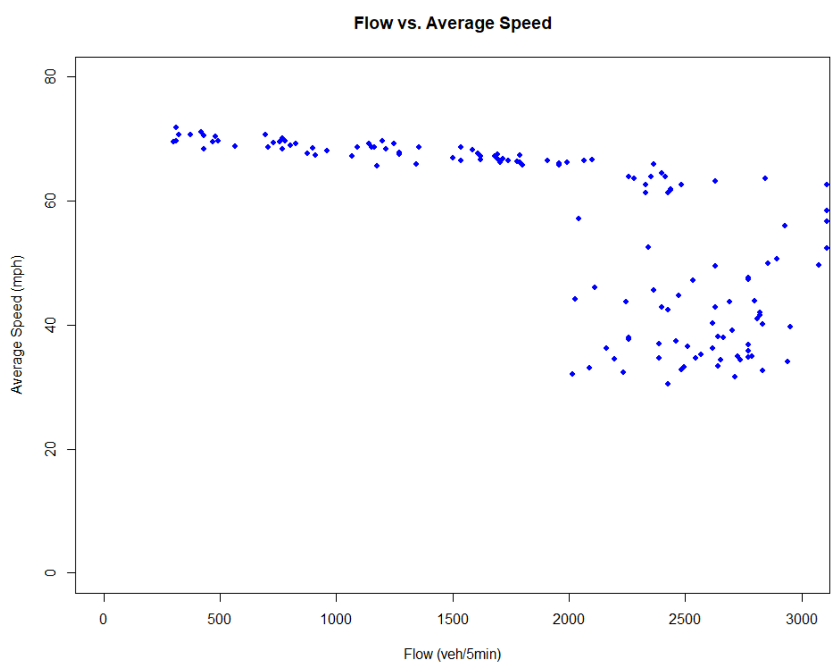


Speed vs Occupancy:

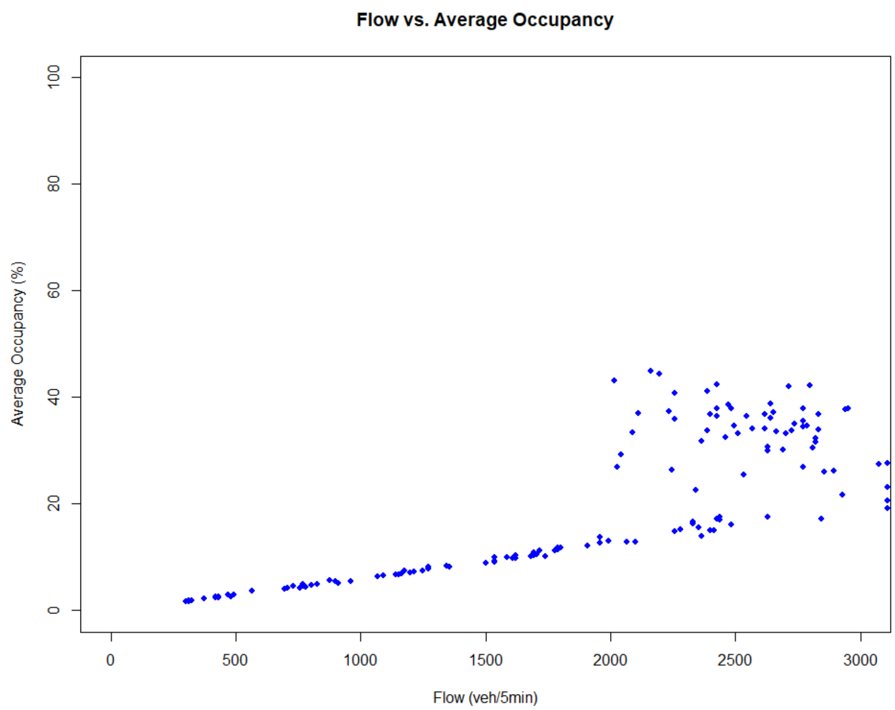


Detector 4:

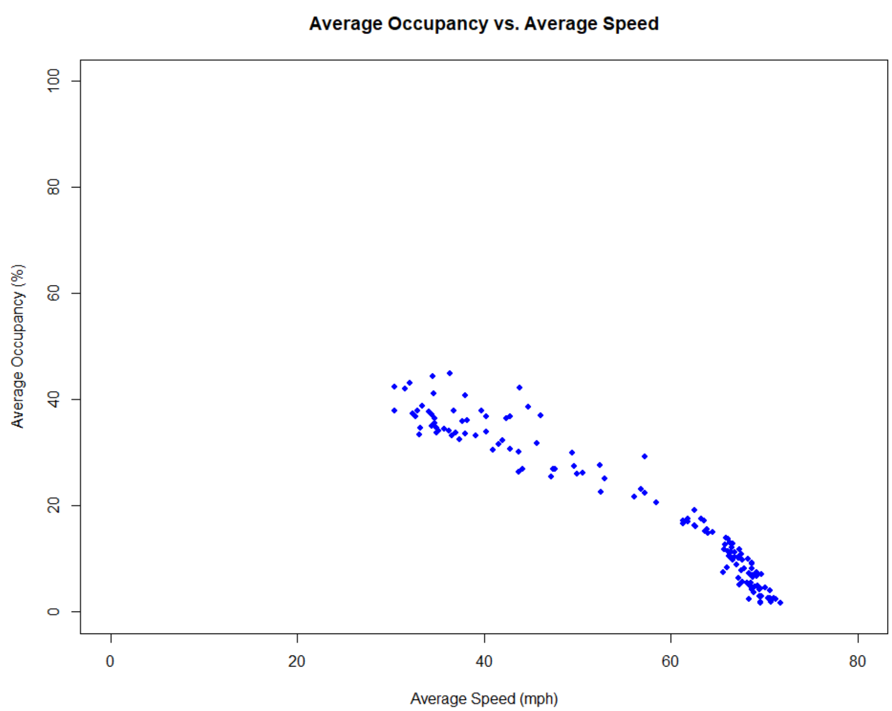
Flow vs Speed:



Flow vs Occupancy:

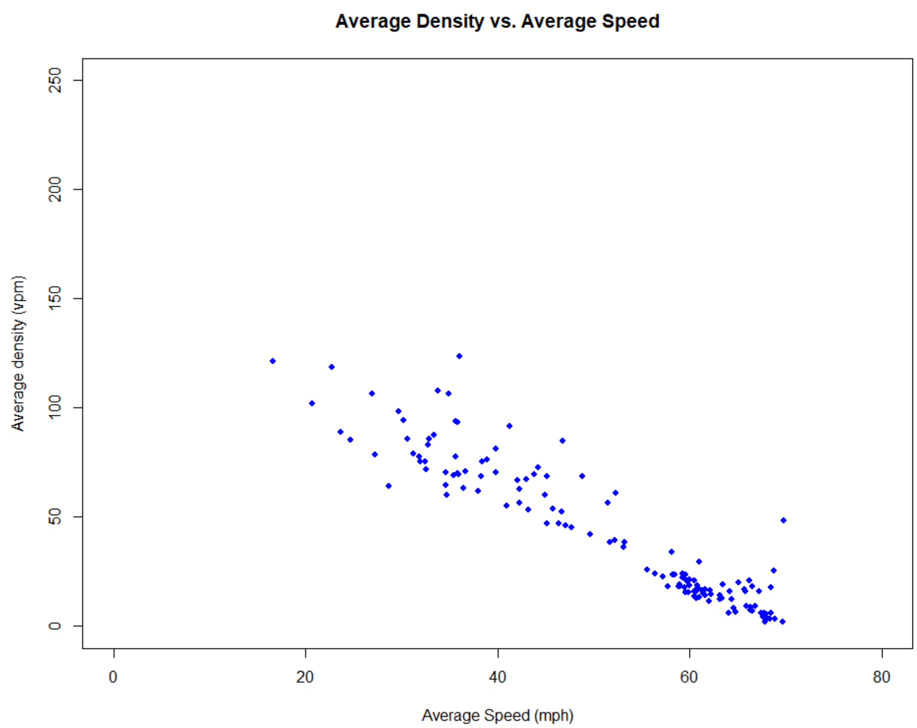


Speed vs Occupancy:

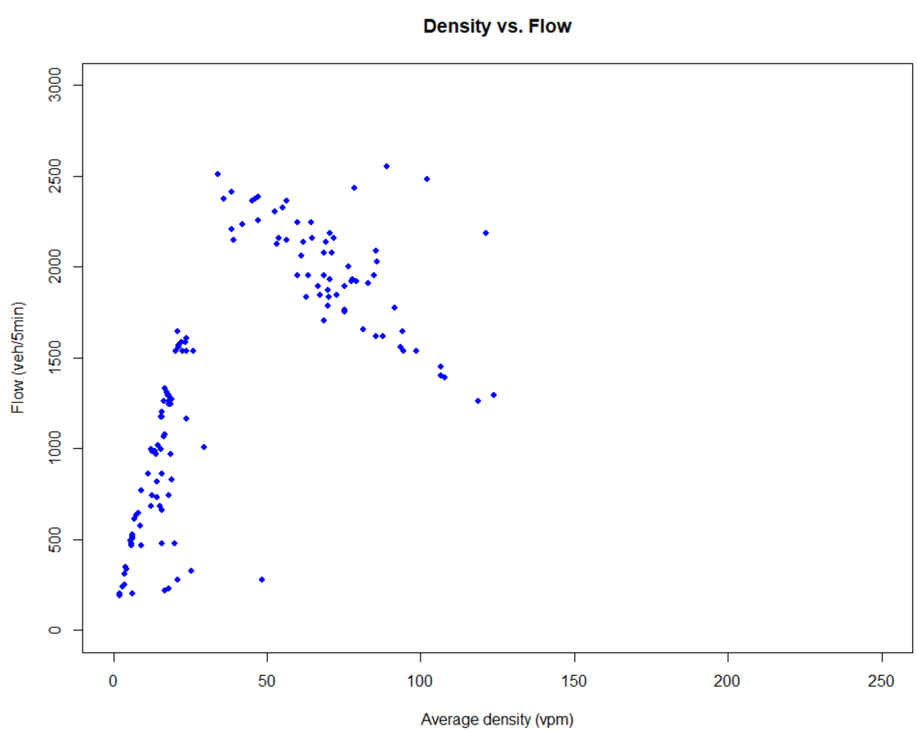


(b) Detector 3:

Speed vs Density:

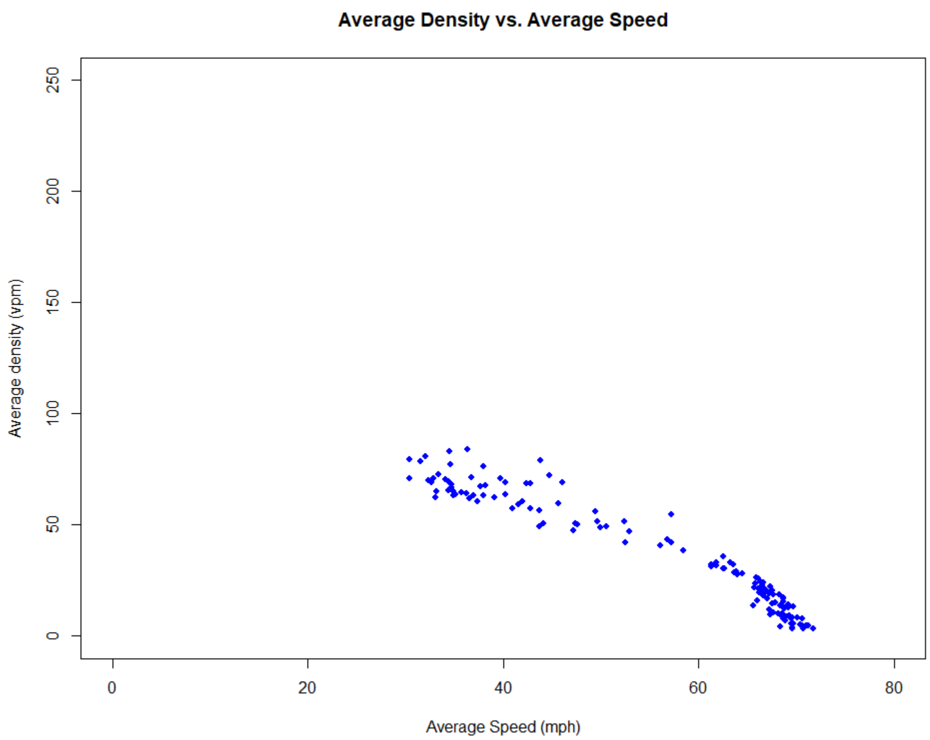


Density vs Flow:

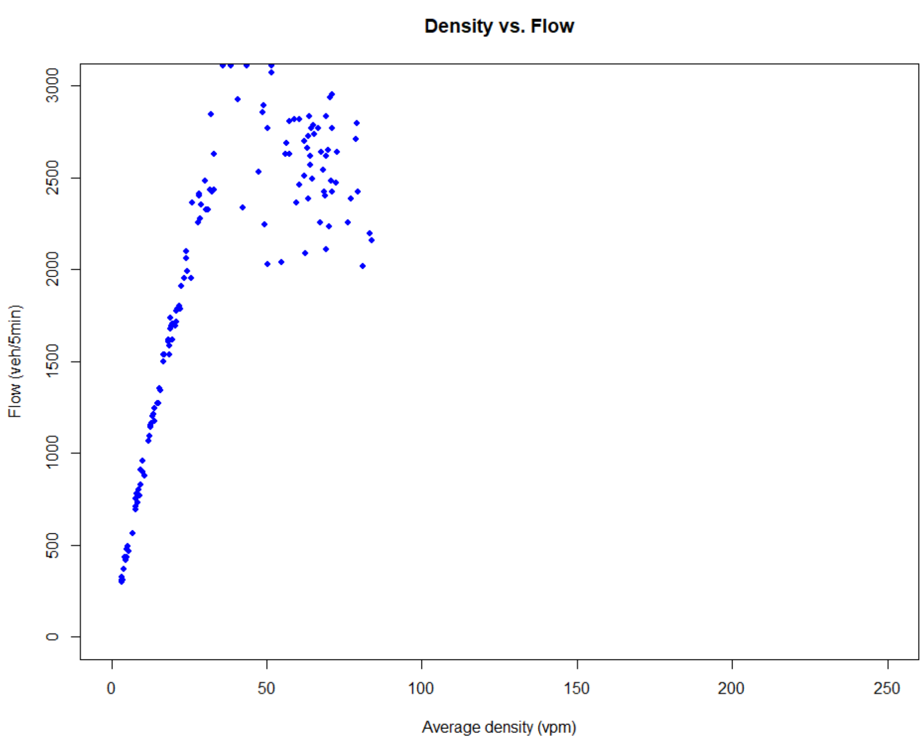


Detector 4:

Speed vs Density:



Density vs Flow:



Differences between detector 3 and 4:

1. For the density vs. speed plot, both detectors demonstrate linear patterns. However, detector 3 have a longer stretch towards lower speed. In addition, the points in detector 3 plot are more scattered than those in detector 4 plot, especially when speed is less than 50 mph.
2. For the density vs flow plot, detector 3 has a longer stretch in the density dimension but shorter stretch in the flow dimension. When flow is under 2000 vehicles / 5 min, detector 4 shows more restricted linear pattern than detector 3 does.

Reason: Detector 4 is located in a road section where the cars travelling by generally have higher speeds and the range of speed captured by detector 4 is more restricted than that captured by detector 3.

(c) From the information given we obtain the following formula:

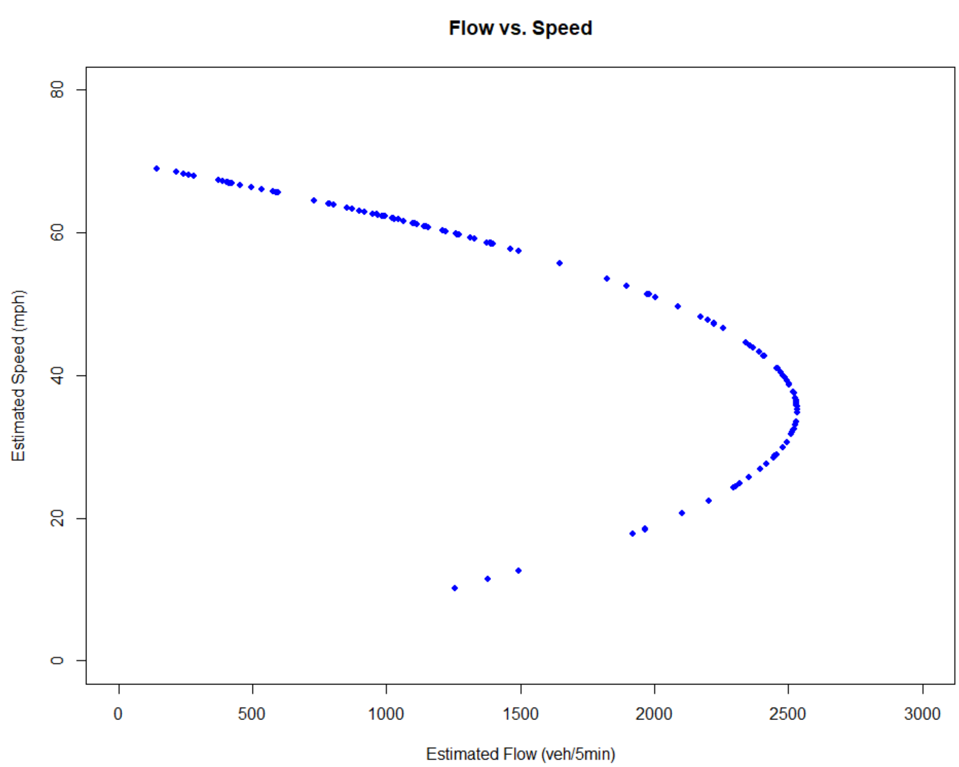
$$k_j = \frac{70k}{70-v}$$

The 90th percentile of kj after dropping all the values greater than 150 is 230.1588 vpm nad this value is treated as the real value of jam density. Then the following formula are used to estimate speed v and flow q:

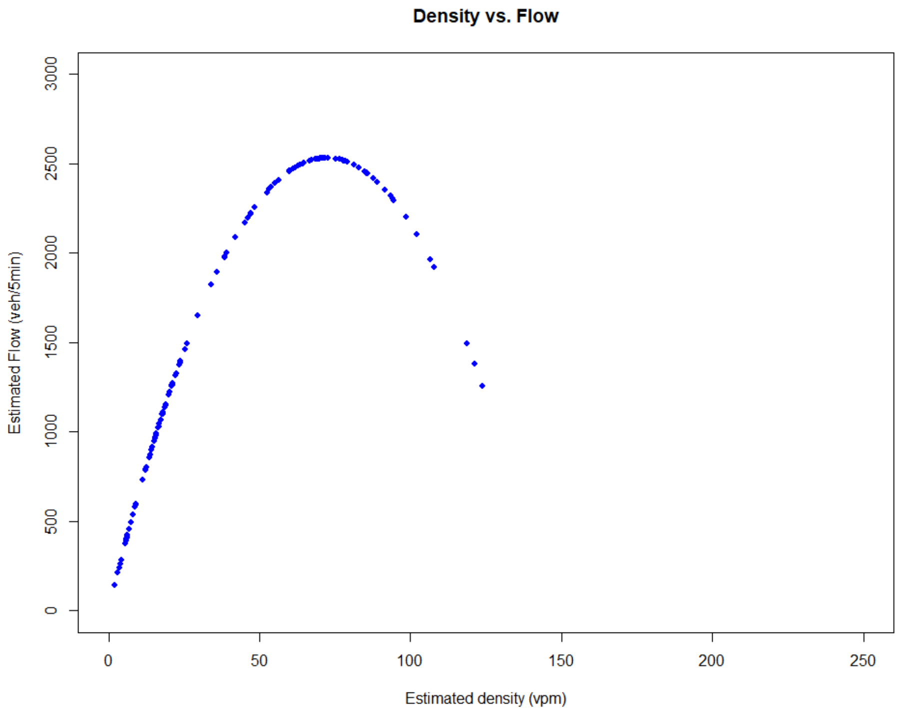
$$V = vf - \frac{vf}{kj} * k$$

$$q = k*v$$

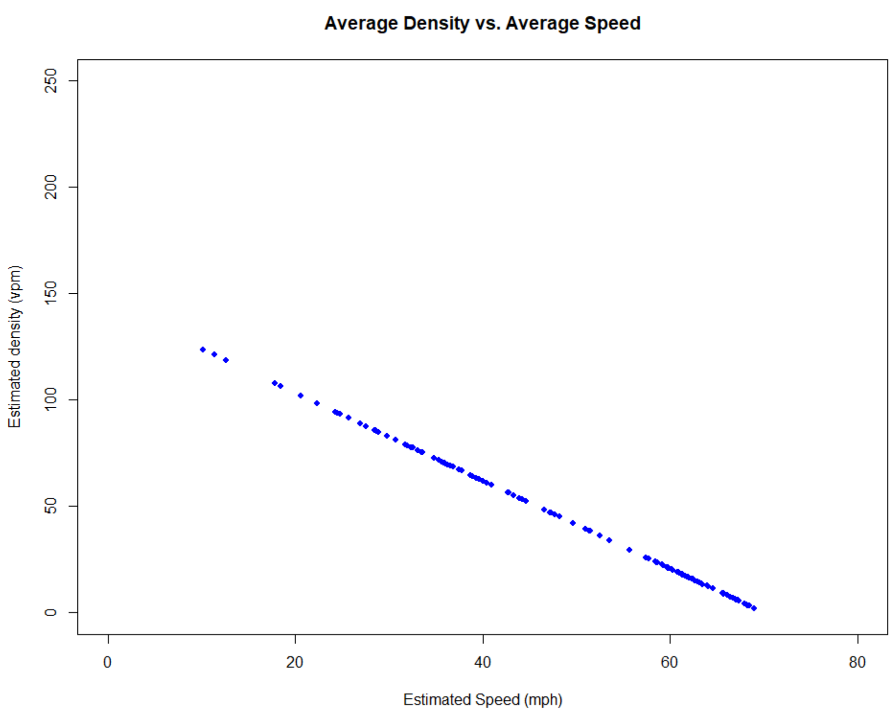
plotting the estimated results:
flow vs speed:



Flow vs Density:

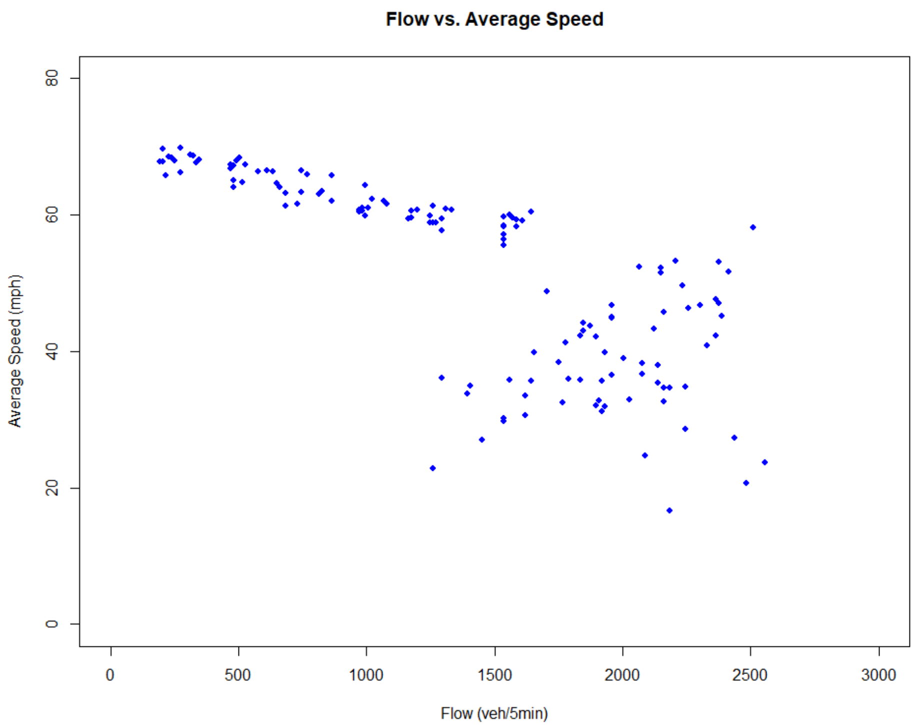


Speed vs Density:

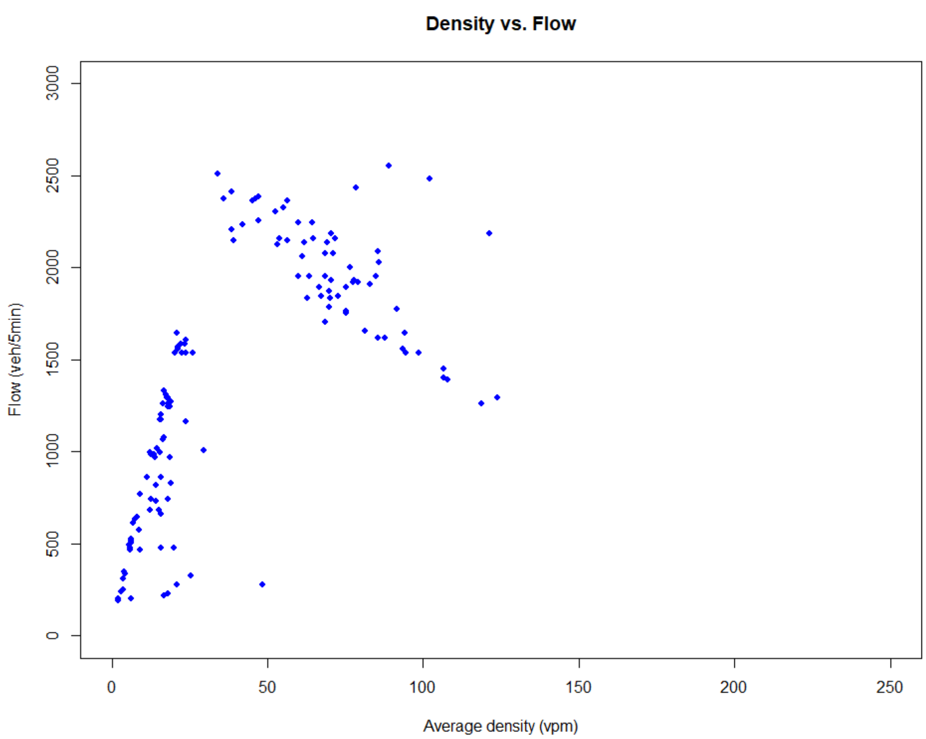


Plotting the observed values:

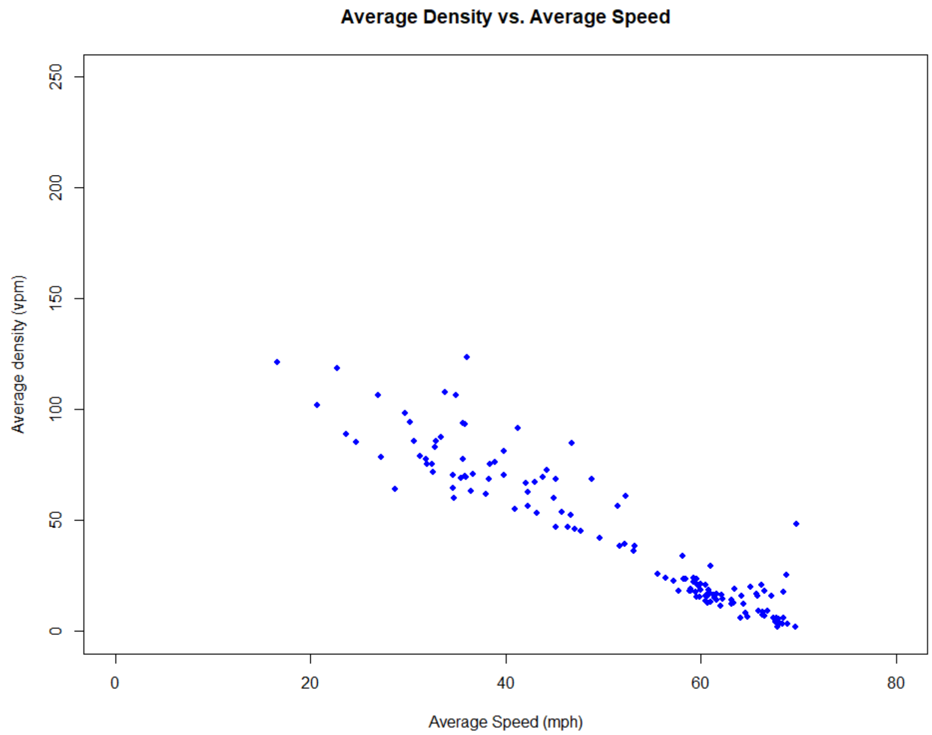
Flow vs speed:



Flow vs density:

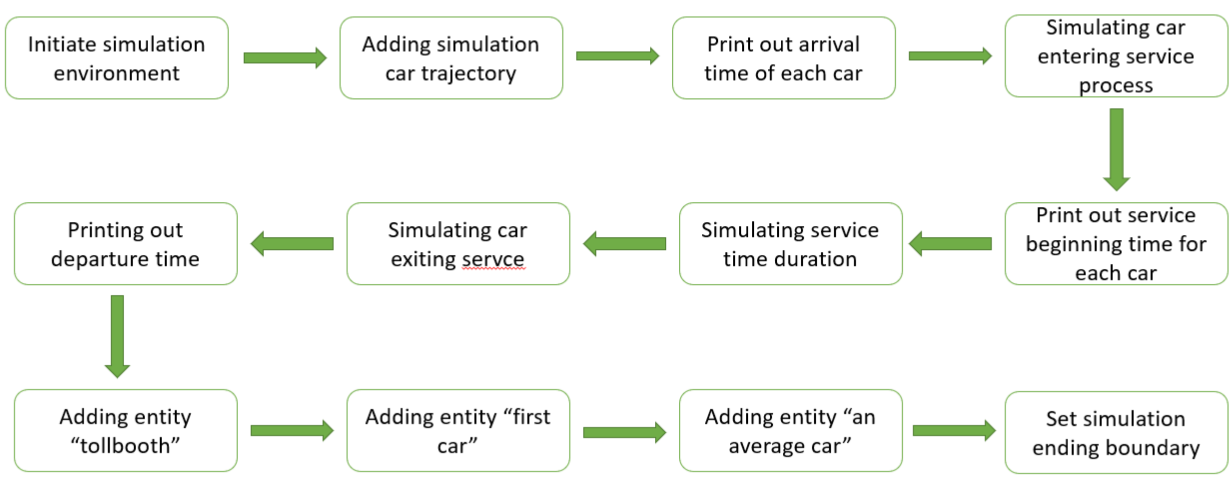


Speed vs density:



Problem 4:

(a) Documentation



Variables:

Car arrival rate: 0.355

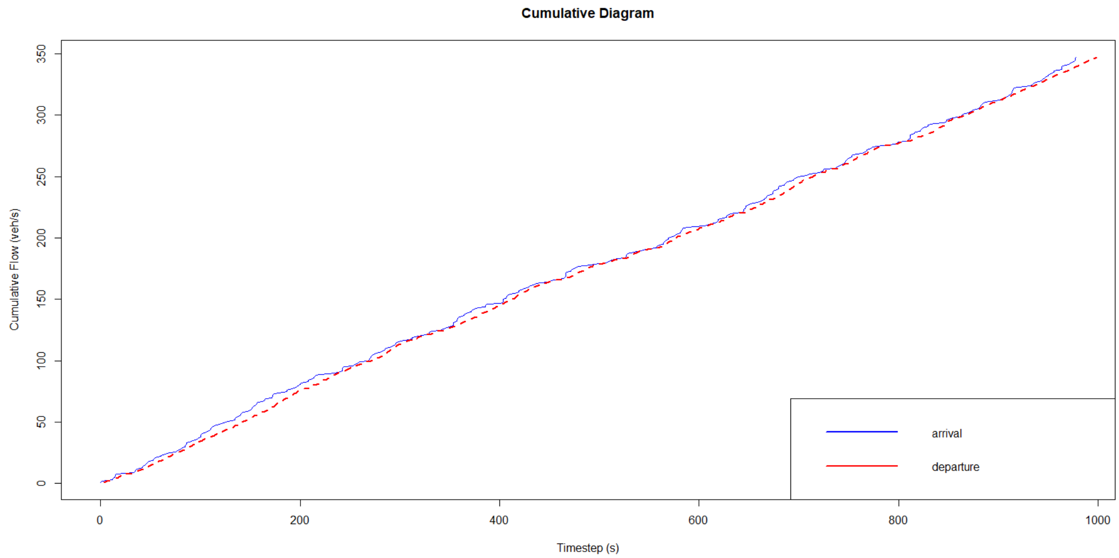
Minimum service time: 1s

Maximum service time: 4s

Simulation time step: 1000s

(b) Average delay: 6.304541 s
Standard deviation: 5.374287 s

(c) Cumulative diagrams



(d) If a random seed is given, the results for 5 different runs are the same.
If no random seed is given, the results are different since the simulation process is a random process. The five results are:
1st run: average delay = 4.028449 , standard deviation = 4.249736 ;
2nd run: average delay = 8.963225, standard deviation = 7.214237;
3rd run: average delay = 5.19175, standard deviation = 5.76098;
4th run: average delay = 9.429984 , standard deviation = 7.91865;
5th run: average delay = 10.05706 , standard deviation = 9.575762.

Problem 5:

(a) When a = 0, b = 2, the true integral value is 1.740;

When a = 0, b = 4, the true integral value is 296.526;

(b) With random seed 1234:
1. n = 100:
when a = 0, b = 2, the MC integral value is 1.488;
when a = 0, b = 4, the MC integral value is 461.578;
2. n = 1000:
when a = 0, b = 2, the MC integral value is 1.766;
when a = 0, b = 4, the MC integral value is 241.337;
3. n = 10000:
when a = 0, b = 2, the MC integral value is 1.7437;
when a = 0, b = 4, the MC integral value is 298.7974;

Without random seed:

1. n = 100:
   when a = 0, b = 2, the MC integral value is 1.663;
   when a = 0, b = 4, the MC integral value is 217.983;
2. n = 1000:
   when a = 0, b = 2, the MC integral value is 1.793;
   when a = 0, b = 4, the MC integral value is 316.607;
3. n = 10000:
   when a = 0, b = 2, the MC integral value is 1.726;
   when a = 0, b = 4, the MC integral value is 292.749;

(c) Comments:

No matter whether a random seed is given or not, for both settings of a and b, the largest n always corresponds to the best simulation of the true results.

The only difference between random seed situation and no random seed situation is that when random seed is absent, the MC simulation result is unstable, i.e., we get different results in different runs. If random seed is provided, we always have the same numbers. However, even though the results are unstable, in no random seed situation the best simulation are always generated by largest n.

Generally speaking, regardless of random seed, both settings of a and b reach the optimal simulation (nearest to the real value) when n = 10000. However, it can be observed from the results that the second setting (a = 0, b = 4) gets better simulation than the first setting.