Bangladesh University of Engineering and Technology, Dhaka

CE 6507

Assignment 1

ON TRAFFIC FLOW MODELING ON FREEWAYS

Submitted in
25 Nov 2024

Submitted by
Arnob Protim Roy
ID : 0424042406
M.Sc Student
Department of Civil Engineering , BUET

Assignment 1

All the steps of my analysis are given below,

(I have used **Lane ID = 2, V_class =2 and location= us-101** for my data**)**
I have first separated the data by vehicle type , lane and highway using python

```python
# Filter the dataset for the specified conditions
filtered_data = data[(data['Lane_ID'] == 2) & (data['v_Class'] == 2) & (data['Location'] == "us-101") & (data['Following'] != 0)]
```

Then using this type of data i have removed the duplicate data as it was mention in the data source website

# Vehicle_ID	Vehicle identification number (ascending by time of entry into section). REPEATS ARE NOT ASSOCIATED.	vehicle_id	Number

Then,
For a row, for a vehicle get the leader vehicle id by Preceding,

# Preceding	Vehicle ID of the lead vehicle in the same lane. A value of '0' represents no preceding vehicle - occurs at the end of the study section and off-ramp due to the fact that only complete trajectories were recorded by this data collection effort (vehicles already in the section at the start of the study period were not recorded).

and then i get the info of the leader vehicle.
So , in the search for a leader vehicle where FrameID is same and Leader_Vehicle_ID= Subject car Preceding id.

Now the column are like this

```
Column names in the dataset:
['Vehicle_ID', 'Frame_ID', 'Total_Frames', 'Global_Time', 'Local_X', 'Local_Y', 'Global_X', 'Global_Y', 'v_length', 'v_Width', 'v_Class', 'v_Vel', 'v_Acc', 'Lane_ID', 'O_Zon
e', 'D_Zone', 'Int_ID', 'Section_ID', 'Direction', 'Movement', 'Preceding', 'Following', 'Space_Headway', 'Time_Headway', 'Location', 'leader_Vehicle_ID', 'leader_Frame_ID',
 'leader_Total_Frames', 'leader_Global_Time', 'leader_Local_X', 'leader_Local_Y', 'leader_Global_X', 'leader_Global_Y', 'leader_v_length', 'leader_v_Width', 'leader_v_Class'
, 'leader_v_Vel', 'leader_v_Acc', 'leader_Lane_ID', 'leader_O_Zone', 'leader_D_Zone', 'leader_Int_ID', 'leader_Section_ID', 'leader_Direction', 'leader_Movement', 'leader_Pr
eceding', 'leader_Following', 'leader_Space_Headway', 'leader_Time_Headway', 'leader_Location']
```

Now I have selected the 500 rows.

Now,
The original equation is this

$$\ddot{x}_{n+1}(t+T) = a_0 \dot{x}_{n+1}^{m}(t+T) \frac{\dot{x}_n(t) - \dot{x}_{n+1}(t)}{(x_n(t) - x_{n+1}(t))^l}$$

It can be written as

$\mathbf{Y = a_0\ X_1^m\ \ X_2 / X_3^l}$

Where
data indicate the filtered dataset.

Y = $\ddot{x}_{n+1}(t+T)$ = `data['v_Acc']`

X_1 = $\dot{x}_{n+1}^{m}(t+T)$ = `data['v_Vel']`

X_2= $\dot{x}_n(t)-\dot{x}_{n+1}(t)$ = `data['leader_v_Vel'] - data['v_Vel']`

X_3= $\left(x_n(t)-x_{n+1}(t)\right)$ = `data['Space_headway']`

Now a new dataset is generated in this format with 500 rows,

Y	X1	X2	X3
-6.48	22	-2.03	42.51
1.28	28.12	-3.12	73.59
2.65	18.77	-2.85	50.86
0	30	-1.5	75.64
0	15	-5.3	52.11
0	0	5	31.95
-0.02	5	1.31	39.41
-1.68	40.08	0.15	129.13
0	20	0	57 .....

(Datasheets are provided separately)

The equation is
$Y = a_0 X_1^m X_2 / X_3^l$

After In Both side,
$\ln(Y) = \ln(a_0) + m \ln(X_1) + \ln(X_2) - l \ln(X_3)$

After curve fitting and regression analysis this dataset using python *[Python code and dataset link are attached in the end*]

this is the a0, m and l of the coefficients where the solution converge in python
a0: 0.2566612097929136, m: 0.4700, l: 0.6347
R-squared: 0.0145 (*Comment: The dataset is not properly following this particular model as R^2 is very low*.)
So

$$\ddot{x}_{n+1}(t+T) = a_0 \dot{x}_{n+1}^{m}(t+T)\frac{\dot{x}_n(t)-\dot{x}_{n+1}(t)}{\left(x_n(t)-x_{n+1}(t)\right)^l}$$

where
a0: 0.2566612097929136, m: 0.4700, l: 0.6347

There is another alternative approach(2) , (as this type of non-linear equation may have many solution)

If,
$$Y = a_0 \, x_1^m \, \frac{x_2}{x_3^l}$$

$$\ln y = \ln a_0 + m \ln x_1 + \ln x_2 - l \ln x_3$$

Let, $\ln y = \underline{Y}$, $\ln a_0 = \underline{a_0}$, $\ln x_1 = x_1$

$\ln x_2 = \underline{x_2}$, $\ln x_3 = \overline{x}_3$, $n = 500$

Now, from regression theory,

$$\Sigma \underline{Y} = \underline{a_0} n + m \Sigma \underline{x_1} + \Sigma \underline{x_2} - l \Sigma \underline{x_3}$$
$$\Sigma \underline{x_1 Y} = \underline{a_0} \Sigma \underline{x_1} + m \Sigma \underline{x_1^2} + \Sigma \underline{x_1 x_2} - l \Sigma \underline{x_1 x_3}$$
$$\Sigma \underline{x_2 Y} = \underline{a_0} \Sigma \underline{x_2} + m \Sigma \underline{x_1 x_2} + \Sigma \underline{x_2^2} - l \Sigma \underline{x_2 x_3}$$

From the updated_dataset_for_excel_regression.csv these values are calculated.

Sum of	ln Y	lnX1	lnx2	lnx3	Sum of	ln(x1)ln(y)	ln(x1)^2	ln(x1)*ln(x2)	ln(x1)*ln(x3)	ln(x2)^2	ln(x2)ln(y)	ln(x2)*ln(x3)
	752.6045	885.1987	752.6125	981.6449		21.912913	30.91940896	21.67489179	36.3538	16.40522	15.63408	25.68475

So, Eventually,

$$866.00 = 500 \underline{a_0} + 950.1 m + 851.6 - 1025.1 l$$
$$1695.7 = 950.14 \underline{a_0} + 1808.8 m + 1614.8 - 1950 l$$
$$1474.9 = 851.6 \underline{a_0} + 1614.8 m + 1464.9 - 1793.4 l$$

The solution matrix,

$$\begin{bmatrix} 500 & 950.1 & -1025.1 \\ 950.1 & 1808.8 & -1950.0 \\ 851.6 & 1614.8 & -1793.4 \end{bmatrix} \begin{bmatrix} \underline{a_0} \\ m \\ l \end{bmatrix} = \begin{bmatrix} 866 - 851.6 \\ 1695.7 - 1614.8 \\ 1474 - 1464. \end{bmatrix}$$

Solving

$$\underline{a_0} = -5.67 \quad \ln a_0 = -5.67, \quad a_0 = e^{-5.67} = 3.44 \times 10^{-3}$$
$$m = 3.01$$
$$l = 1.66$$

This can be another set of solution.

Third Approach

Third approach,

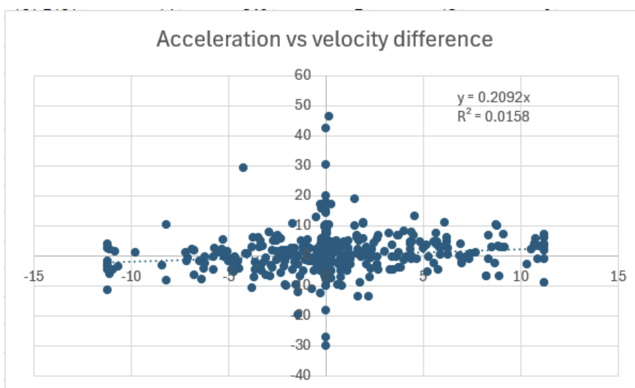
To get linear CFM,

taking $l=0$, $m=0$

$$\ddot{x}_{n+1}(t+T) = a_0(\dot{x}_n(t) - \dot{x}_{n+2}(t))$$

Plotting,

$$\ddot{x}_{n+1}(t+T) \text{ vs } (\dot{x}_n(t) - \dot{x}_{n+2}(t))$$



$$\ddot{x}_{n+1}(t+T) \text{ vs } (\dot{x}_n(t) - \dot{x}_{n+2}(t))$$

$$y = 0.209x$$

Comparing, $a_0 = 0.209$;

Now,

Let $\dot{x}_{n+1}(t+T) = y$

and, $x_n(t) - x_{n+1}(t) = x$

So, $\frac{\partial y}{\partial x} = \ddot{x}_{n+1}(t+T)$

placing the value on original Equation,

$$\frac{\partial y}{\partial x} = a_0 \frac{x^m}{x^l} \frac{\partial x}{\partial t}$$

$$\Rightarrow y^{-m} dy = a_0 x^{-l} dx$$

$$\Rightarrow \int y^{-m} dy = \int a_0 x^{-l} dx$$

$$\Rightarrow \frac{y^{-m+1}}{-m+1} = a_0 \frac{x^{-l+1}}{-l+1}$$

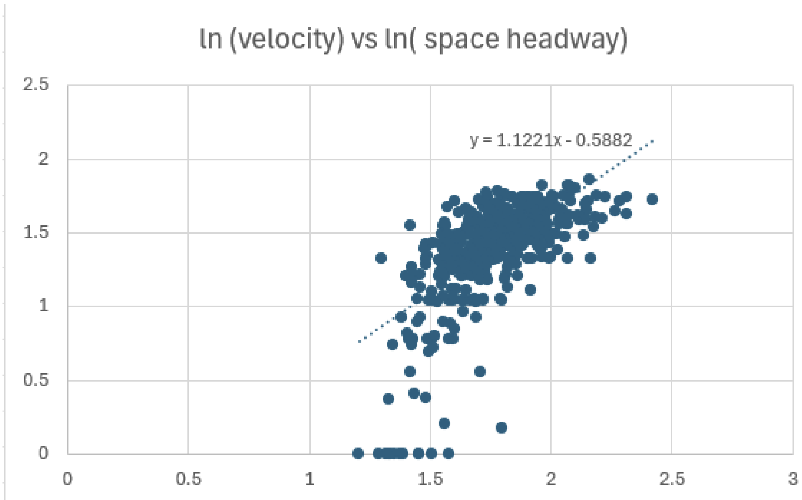
$$\Rightarrow y^{-m+1} = a_0 \frac{-m+1}{(-l+1)} x^{-l+1}$$

$$\Rightarrow \ln Y = \frac{1}{-m+1} \ln\left(\frac{a_0(-m+1)}{-l+1}\right) + \frac{(-l+1)}{-m+1} \ln X$$

Now plotting,

$\ln y$ vs $\ln x$

or $\ln(\dot{x}_{n+1})$ vs $\ln(\text{spaceheadway})$



$$\ln Y = \frac{1}{-m+1} \ln\left(a_0 \frac{(-m+1)}{-l+1}\right) + \frac{(-l+1)}{-m+1} \ln X$$

$$\ln Y = -0.5882 + 1.12 \ln X$$

Comparing,

$$\frac{1-l}{1-m} = 1.12, \quad l-1 = 1.12(m-1)$$

And,

$$\frac{1}{-m+1} \ln\left(a_0 \frac{-m+1}{-l+1}\right) = -0.588$$

$$\Rightarrow \frac{1}{1-m} \ln\left(\frac{0.209}{1.12}\right) = -0.588$$

another set of solution

$$\left\{ \begin{array}{l} m = -1.85, \quad l = -2.192 \\ a = 0.209 \end{array} \right.$$

Ans no 2

We know, density = 1/ distance_headway

Now I will curve fitting this equation (Using Python Programming)

```
Y  = a  ln (b / x)   ( Greenberg model)
Where, Y  = Speed , X = Density
```

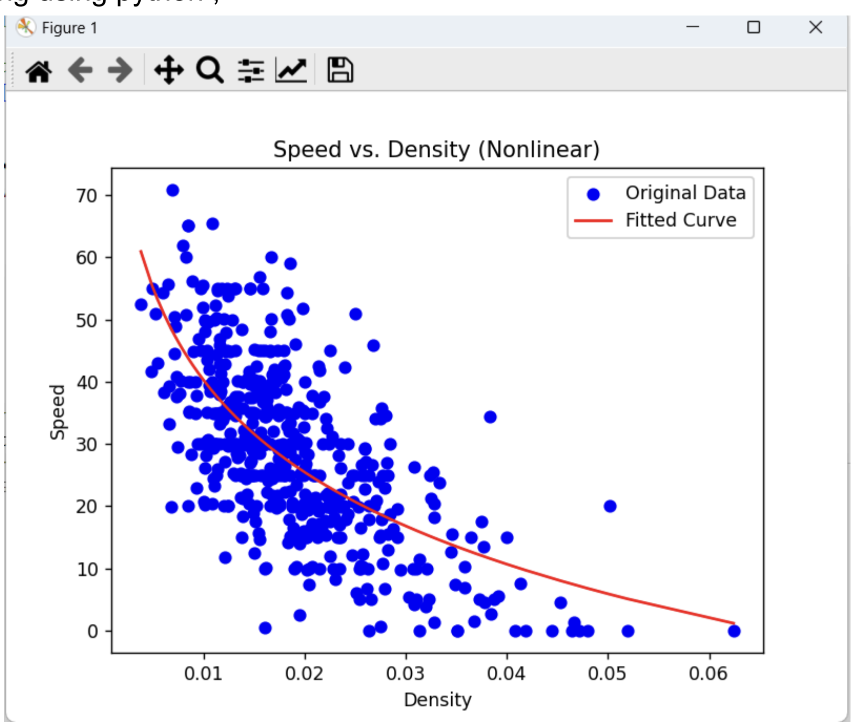
Analysis-
Model: a = 21.1389841697585 ,
b = 0.0632646411137779
R-squared: 0.45199910607969596

Nonlinear Model: a = 21.31823879833061 ,b = 0.06599902025673599
R-squared: 0.4345099030151789

Y = 21.31 ln (0.06/ X)

(Also b can be told as jam density)

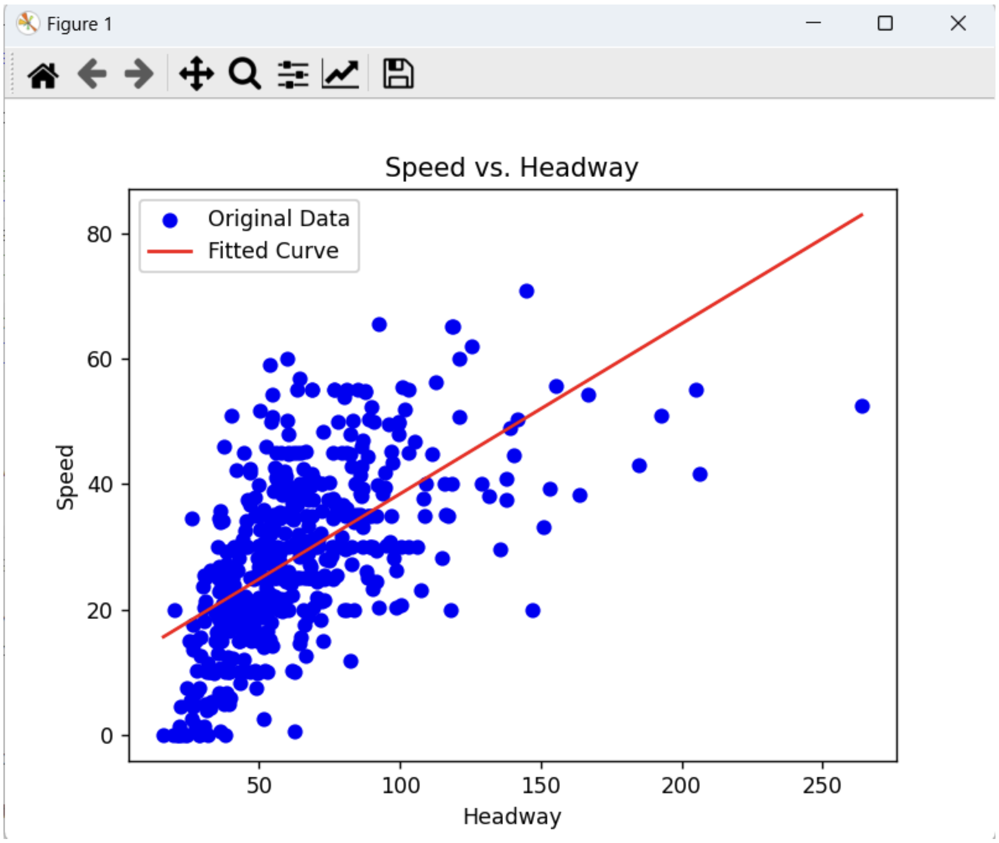
After Curve fitting using python ,



Now if I want Speed vs. Space headway model,
the relationship can be considered as linear .

The simple relationship is
Y = a X + B
Speed = a * distance_headway + b

the curve is generated using python plot,



Analysis :
Model: a = 0.2713382282717081 ,b = 11.298975329567456
R-squared: 0.3461301719869715

So , Speed = 0.27 * distance_headway + 11.99

Note:

- For the first part, Three different approaches are provided.
- the units are fixed as they are provided. units are chosen very carefully after studying all the documentation.
- Analysis are basically done in python (Except the 2nd and 3rd alternative analysis using regression technique, which has been done in excel , but the first approach was done by python). All python codes and dataset are uploaded at this link (https://github.com/trewto/CE-6507-Assignment-/tree/main/Assignment%201)
- 1. combined_dataset_500_with_leader.csv here is to 500 dataset where leader vehicles are merged ensuring same frame
- 1b) updated_combine_dataset X1X2X3.csv : this is used for the 1st approach
- 2. excelworkfor_regression.csv ; this is use for second apporach for the first problem
- 3. combined_dataset_for_CFM_3rdmethod.csv, this is used for 3rd metheod