École Polytechnique fédérale de Lausanne

CIVIL-557

Project - Modelling

Jingren Tang

1 Task 1

1. As shown in Figure 1, the model with a lag time of 1 exhibits the smallest Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values, indicating the best model fit among the candidates. Additionally, it has the highest adjusted Rho square and log likelihood values, further confirming the superior performance of the model with lag time 1. Therefore, we choose the lag time to be 1 for our analysis.

Model	Final Log Likelihood	Adjusted Rho Squared	AIC	BIC
lag1	-18212.25013	0.12194469	36444.50027	36520.88376
lag2	-19770.97524	0.04683613	39561.95049	39638.33398
lag3	-20297.00239	0.02148905	40614.00479	40690.38828
nolag	-20117.17605	0.03015414	40254.3521	40330.73559

Table 1: Comparison of Different Lag Models

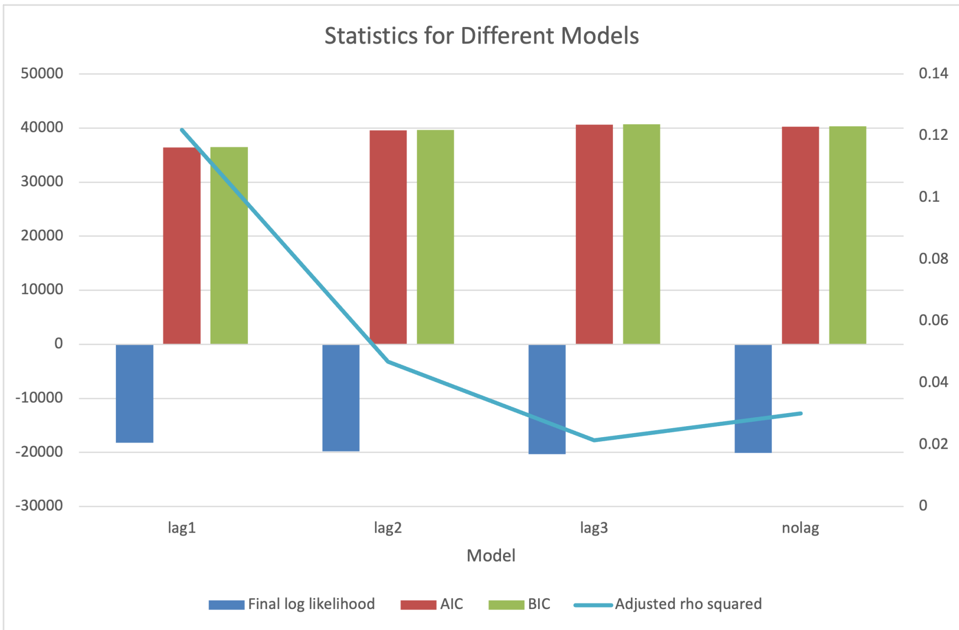


Figure 1: Model comparison results for different lag times.

2. The result shows that the reaction time of 1 second significantly improves the model fit, whereas the reaction times of 2 and 3 seconds do not show a notable improvement in the fitting performance.The superior performance of the 1-second lag model suggests that the reaction time of drivers in the analyzed data set is approximately 1 second. This means that, on average, drivers take about 1 second to respond to changes in the speed of the car they are following.

2 Task 2

Sign, Magnitude, and Significance of Parameters

1. **Sign**:

- **Positive Parameters**:
  - α_{acc} (1.894), β_{acc} (0.37), γ_{acc} (0.737), λ_{acc} (0.926)
  - γ_{dec} (0.753), λ_{dec} (0.923)
- **Negative Parameters**:
  - α_{dec} (-2.715), σ_{acc} (-0.248), σ_{dec} (-0.216)

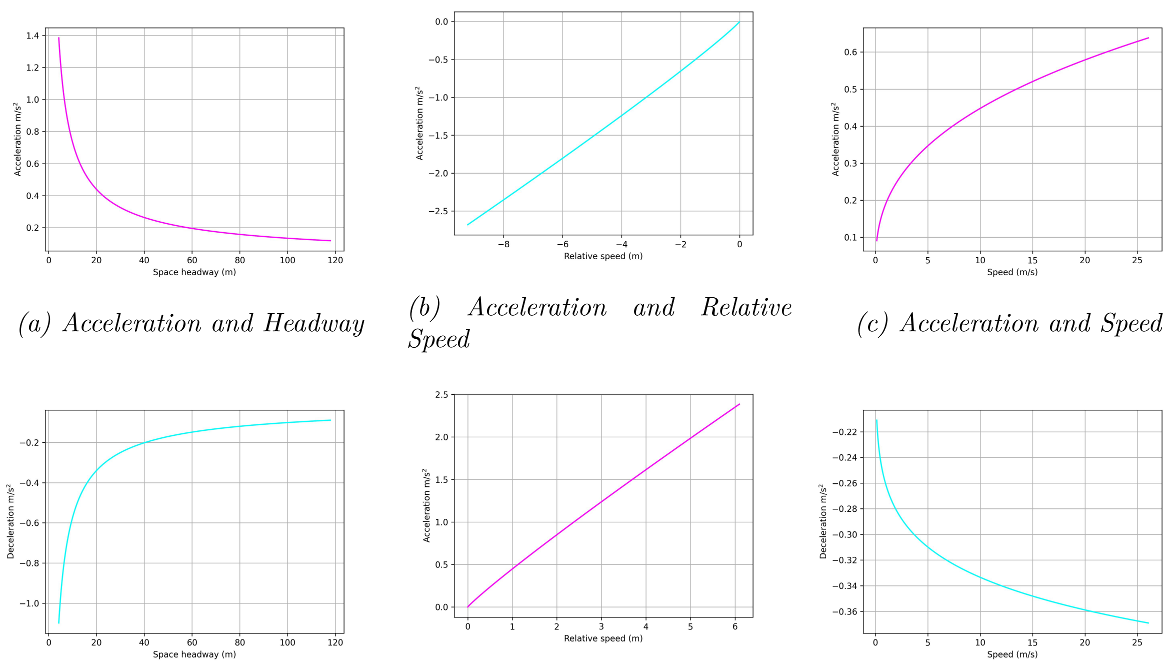
- β_{dec} (0.106) is positive, but its significance varies

2. **Magnitude**:

- λ_{acc} (0.926) and λ_{dec} (0.923) have large positive values, indicating strong positive effects.
- α_{dec} (-2.715) has a large negative value, indicating a strong negative effect.

3. **Significance**:

- All parameters have p-values close to 0.0, indicating that they are statistically significant at conventional significance levels (0.05).
- Robust standard errors confirm the significance of most parameters except for β_{dec}, where the robust p-value is 0.052, slightly above the 0.05 threshold.



(a) Acceleration and Headway

(b) Acceleration and Relative Speed

(c) Acceleration and Speed

(d) Deceleration and Headway

(e) Deceleration and Relative Speed

(f) Deceleration and Speed

Figure 2: Comparison of Acceleration and Deceleration Parameters

Range and Effect of Explanatory Variables

For headway and speed, the influence of the explanatory variables has a much larger effect on acceleration or deceleration when their values are relatively small. In contrast, for relative speed, the relationship with acceleration or deceleration appears to be linear regardless of whether the explanatory variable value is large or small.

Interpretation of Parameters in Practice

As the headway increases, acceleration decreases rapidly because a larger headway allows drivers to avoid significant acceleration or deceleration. Acceleration exhibits an almost linear relationship with relative speed, indicating that as relative speed increases, acceleration also increases proportionally. At low speeds, acceleration decreases rapidly because drivers want to quickly accelerate to a comfortable speed. The behavior of deceleration is similar: as headway increases, deceleration decreases. Deceleration also shows an almost linear relationship with relative speed. However, at low speeds, the rate of decrease in deceleration is not as pronounced as that of acceleration, as drivers are more inclined to accelerate quickly when at low speeds.

3 Task 3

Sign, Magnitude, and Significance of Parameters

1. **Sign**:

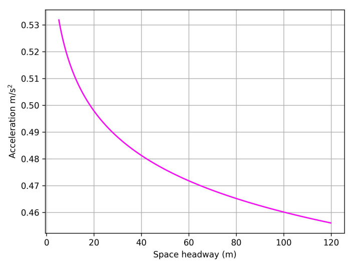
   - **Positive Parameters**:
     - α_{acc} (0.627), γ_{dec} (0.645), λ_{acc} (0.578), λ_{dec} (0.595), σ_{acc} (0.373), σ_{dec} (0.403)
   - **Negative Parameters**:
     - α_{dec} (-6.147), β_{acc} (-0.045), β_{dec} (-0.096)

2. **Magnitude**:

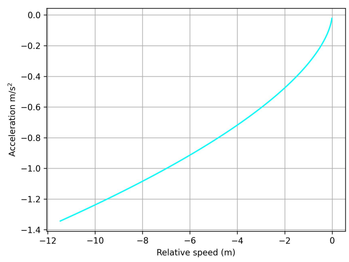
   - α_{dec} (-6.147) has a large negative value, indicating a strong negative effect.
   - λ_{acc} (0.578) and λ_{dec} (0.595) have large positive values, indicating strong positive effects.
   - σ_{acc} (0.373) and σ_{dec} (0.403) also have significant positive values.

3. **Significance**:

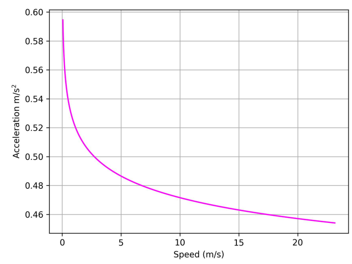
   - Most parameters have p-values close to 0.0, indicating that they are statistically significant at conventional significance levels (0.05).
   - However, β_{acc}, β_{dec}, and γ_{acc} have high p-values (0.562, 0.197, and 0.658, respectively), indicating that they are not statistically significant. These parameters should be interpreted with caution or potentially excluded from the model.



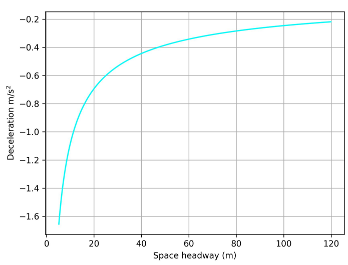
(a) Acceleration and Headway



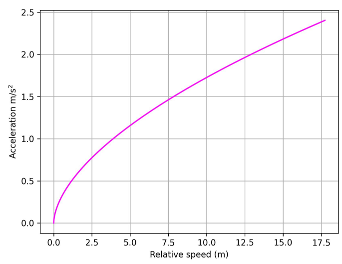
(b) Acceleration and Relative Speed



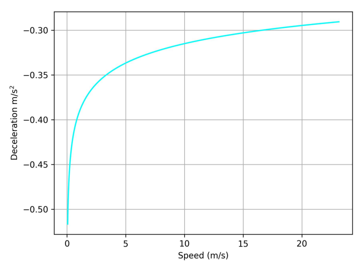
(c) Acceleration and Speed



(d) Deceleration and Headway



(e) Deceleration and Relative Speed



(f) Deceleration and Speed

Figure 3: Comparison of Acceleration and Deceleration Parameters

Range and Effect of Explanatory Variables

For headway, the influence of the explanatory variables has a much larger effect on acceleration or deceleration when their values are relatively small. In contrast, for relative speed, the relationship with acceleration or deceleration appears to be linear regardless of whether the explanatory variable value is large or small. The plot of speed versus acceleration or deceleration appears counterintuitive, and the associated parameters have very high p-values, indicating they are not statistically significant. Therefore, we can disregard these parameters and the plot, as it may be incorrect.

Interpretation of Parameters in Practice

As the headway increases, acceleration decreases rapidly because a larger headway allows drivers to avoid significant acceleration or deceleration. The convexity of relative speed differs for acceleration and deceleration. This is reasonable because when the relative speed is sufficiently high, drivers feel safe to accelerate more, resulting in larger acceleration. Similarly, the same applies to deceleration; as the relative speed gets close to 0, the vehicle may decelerate more significantly to ensure safety.

4 Task 4

We use t-test of individual parameter equivalence to distinguish the difference of two models. As table 2 shows, most of the parameters have higher t value than 1.96 or lower t value than -1.96. Therefore, we conclude that many parameters are significantly different and we should not directly use the model from location 2 to estimate that for location 1.

Parameter	**Value**
alpha_acc	3.164475
alpha_dec	1.075693
beta_acc	3.849695
beta_dec	2.326300
gamma_acc	4.850727
gamma_dec	0.672561
lamda_acc	5.648279
lamda_dec	4.563927
sigma_acc	-39.755451
sigma_dec	-34.766504

Table 2: Parameter Values

5 Task 5

5.1 Sign, Magnitude, and Significance of Parameters

1. **Sign**:

- **Positive Parameters**:
  - $\tau_{mean}(-0.408)$
  - $\alpha 1(0.348)$
  - $\alpha 2(0.007)$
  - $\beta 2(2.265)$
- **Negative Parameters**:
  - $\sigma_\tau(-1.356)$
  - $\beta 1(-0.606)$

– $\sigma(-0.237)$

2. **Magnitude**:

- $\beta2(2.265)$ and $\alpha1(0.348)$ have large positive values, indicating strong positive effects.
- $\sigma_\tau(-1.356)$ and $\sigma(-0.237)$ have large negative values, indicating strong negative effects.

3. **Significance**:

- All parameters except for $\beta1$ have p-values close to 0.0, indicating that they are statistically significant at conventional significance levels (0.05)

Parameter	Estimate	s.e.	t-ratio0	p-value	Rob s.e.	Rob t-ratio0	Rob p-value
tau_mean	-0.408	0.042	-9.782	0.0	0.055	-7.438	0.0
sigma_tau	-1.356	0.102	-13.277	0.0	0.102	-13.227	0.0
alpha1	0.348	0.005	67.816	0.0	0.012	28.873	0.0
alpha2	0.007	0.001	8.948	0.0	0.002	3.306	0.001
beta1	-0.606	2.256	-0.269	0.788	3.447	-0.176	0.86
beta2	2.265	0.231	9.807	0.0	0.4	5.661	0.0
sigma	-0.237	0.006	-40.947	0.0	0.011	-21.693	0.0

Table 3: Parameter Estimates and Statistical Values

Given that all relationships between the explanatory variables and the dependent variable are linear, the effect should be consistent across the entire range. Positive parameters (e.g., $\alpha1$ and $\beta2$) suggest they contribute to higher acceleration, while negative parameters (e.g.,$\beta1$) indicate the opposite effect.

5.2 Issues with the specification of the desired headway

The desired headway may be estimated as a negative value, which is not reasonable. To solve this, we can transform it into an exponential form:

$$\Delta\bar{X}_n(t) = e^{\beta_1+\beta_2 V_n(t-\tau)}$$

This ensures the desired headway is always positive, making the model more accurate and realistic.

6 Task 6

6.1 Does adding distance significantly improve model fit? Justify your answer.

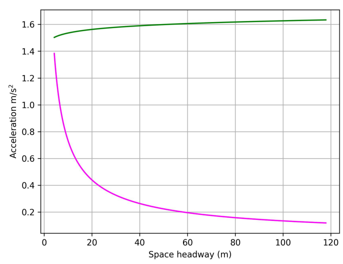
In the given data:

- **Before adding space headway:**
  - Final log likelihood: -4995.716788032517
  - AIC: 10023.433576065034
  - BIC: 10145.647167177342
- **After adding space headway:**
  - Final log likelihood: -5003.393806371175

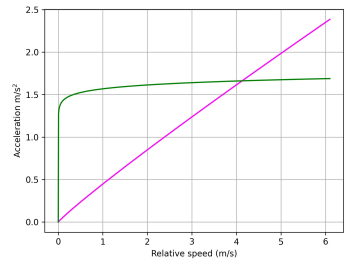
– AIC: 10038.78761274235
– BIC: 10161.001203854657

After adding space headway, both AIC and BIC values increased, indicating a worse model fit. The final log likelihood also decreased. Therefore, adding distance does not significantly improve the model fit; in fact, it worsens it.

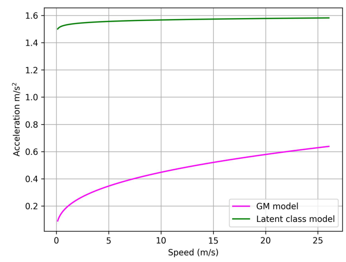
6.2 Discuss the sign, magnitude, and significance of parameters.



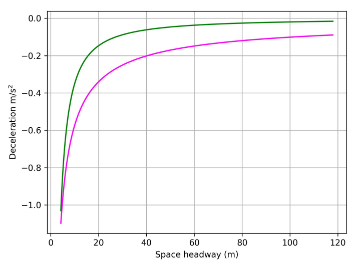
(a) Acceleration and Headway



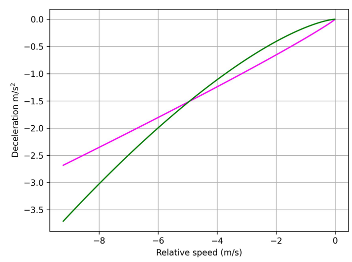
(b) Acceleration and Relative Speed (Positive)



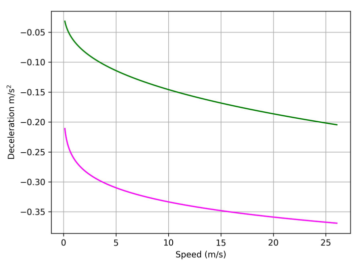
(c) Acceleration and Speed



(d) Deceleration and Headway



(e) Deceleration and Relative Speed (Negative)



(f) Deceleration and Speed

Figure 4: Comparison of Acceleration and Deceleration Parameters

6.3 Compare the impact of explanatory variables

For the latent class model, the acceleration remains around 1.5 regardless of the variation in space headway, relative speed, or speed. This is reasonable because drivers may try to accelerate as long as they perceive it as safe. However, for the deceleration part, the behavior of the GM model follows the same trend as the latent class model, although the deceleration values are always lower than those of the GM model. This may be because they do not need to decelerate as much when they are confident it is safe.

6.4 How do probabilities (of acceleration, deceleration, and do-nothing state) change?

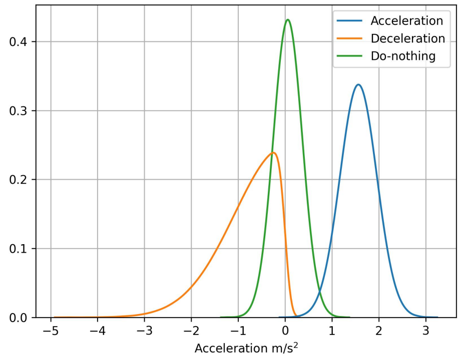


Figure 5: Densities

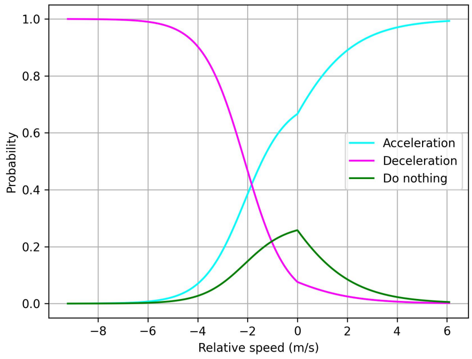


Figure 6: State Probability

Figure 7: Comparison of Densities and State Probability

It can be seen that the deceleration distribution is right-skewed, while the distributions for doing nothing and accelerating are normal distributions. This indicates that people tend to choose a slightly smaller deceleration rate, possibly because it feels better while driving. For doing nothing or accelerating, they control the values within a relatively normal distribution with a target acceleration.

When the relative speed is negative, people are more likely to decelerate, while when the relative speed is higher, they are more likely to accelerate. Naturally, the probability of doing nothing is more densely distributed around zero relative speed.

6.5 Would you recommend the latent class model over the GM model?

Metric	Latent Model	GM Model
Log Likelihood at Zeros	-12961.23386546	-20752.96382237
Initial Log Likelihood	-12961.23386546	-20752.96382237
Final Log Likelihood	-4995.716788032517	-18212.250134826583
Rho Squared	0.61456472	0.12242655
Adjusted Rho Squared	0.61333027	0.12194469
AIC	10023.433576065034	36444.50026965317
BIC	10145.647167177342	36520.88376409836

Table 4: Comparison of Latent and GM Models

Yes. The latent class model has a significantly higher final log likelihood. The latent class model has much higher rho squared and adjusted rho squared values, suggesting it explains the variance in the data better. The latent class model has substantially lower AIC and BIC values, indicating it achieves a better balance between model complexity and fit.

AIC and BIC. AIC/BIC balances model fit and complexity by penalizing the number of parameters. A lower AIC/BIC indicates a model that fits the data well while avoiding overfitting. It is suitable for comparing models that are not necessarily nested.