Exploratory Data Analysis on los_angeles_10.csv

USC Machine Learning Team

11/2/2017

Data Loading

```
# load data file
if (!exists("LA10.raw")) {
    LA10.raw <- read.csv("./los_angeles_10.csv")
}
print(nrow(LA10.raw))
```

```
## [1] 2119
```

Data Cleaning & Exploratory Analysis

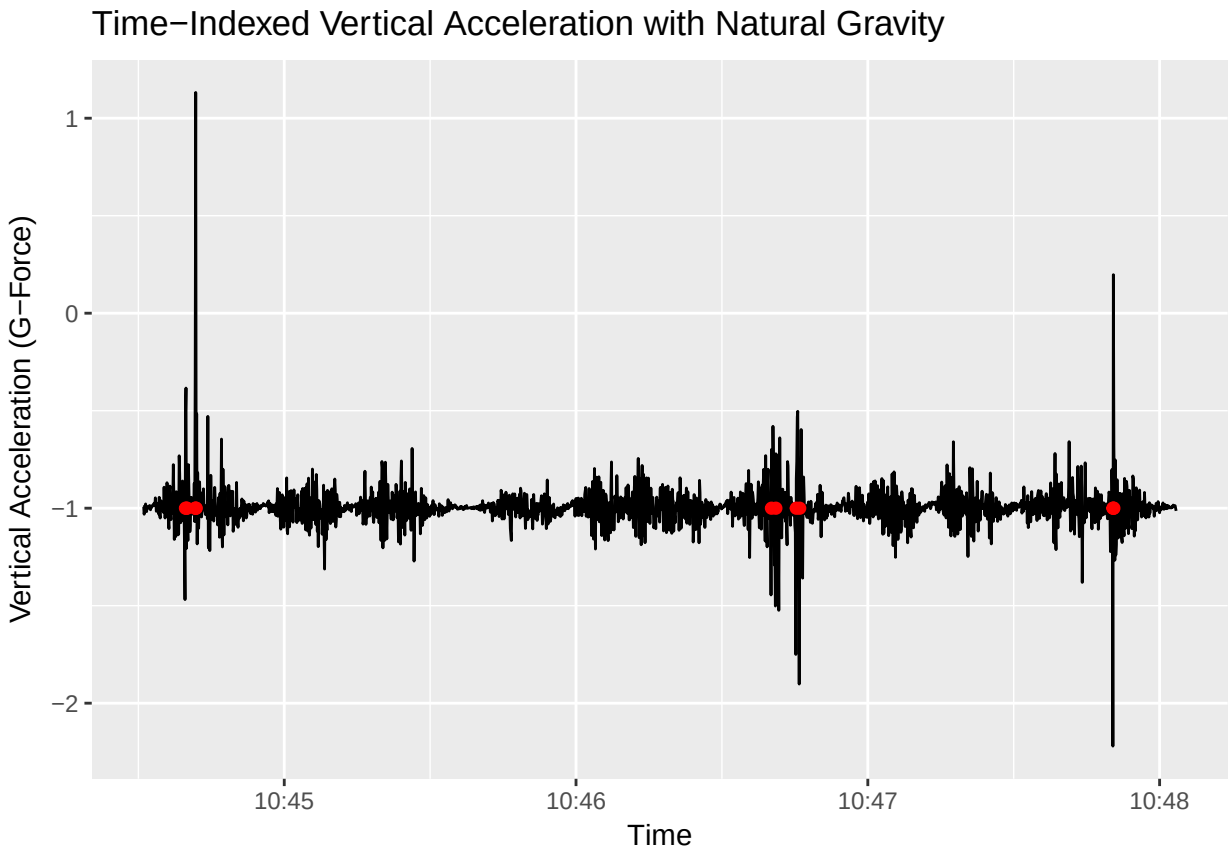
Data Cleaning & Exploratory Analysis are iterative, starting at next page.

Epoch 1: Display of Vertical Acceleration with Natural Gravity

Date: the time stamp of a data point
Speed: traveling speed of the vehicle
forw_accel: forward acceleration (front and back) of the vehicle
hori_accel: horizontal acceleration (left and right) of the vehicle
vert_accel_G: vertical acceleration (up and down) of the vehicle with natural gravity (-1G)

```
LA10.valid = LA10.raw[, c("Date", "Latitude", "Longitude", "Speed")]
# validate DateTime format
LA10.valid$Date <- as.POSIXct(LA10.valid$Date, format="%Y-%m-%d %H:%M:%OS")
# specify orientation of accelration
LA10.valid$forw_accel = LA10.raw$X
LA10.valid$hori_accel = LA10.raw$Y
LA10.valid$vert_accel_G = LA10.raw$Z
# mark speed bumps
for (i in 1:nrow(LA10.valid)) {
    LA10.valid$speedbump[i] = "no"
}
for (i in c(88, 90, 91, 108, 109, 110, 1287, 1296, 1338, 1345, 1989, 1990, 1991)) {
    LA10.valid$speedbump[i] = "yes"
}
```

```
## Loading required package: ggplot2
```

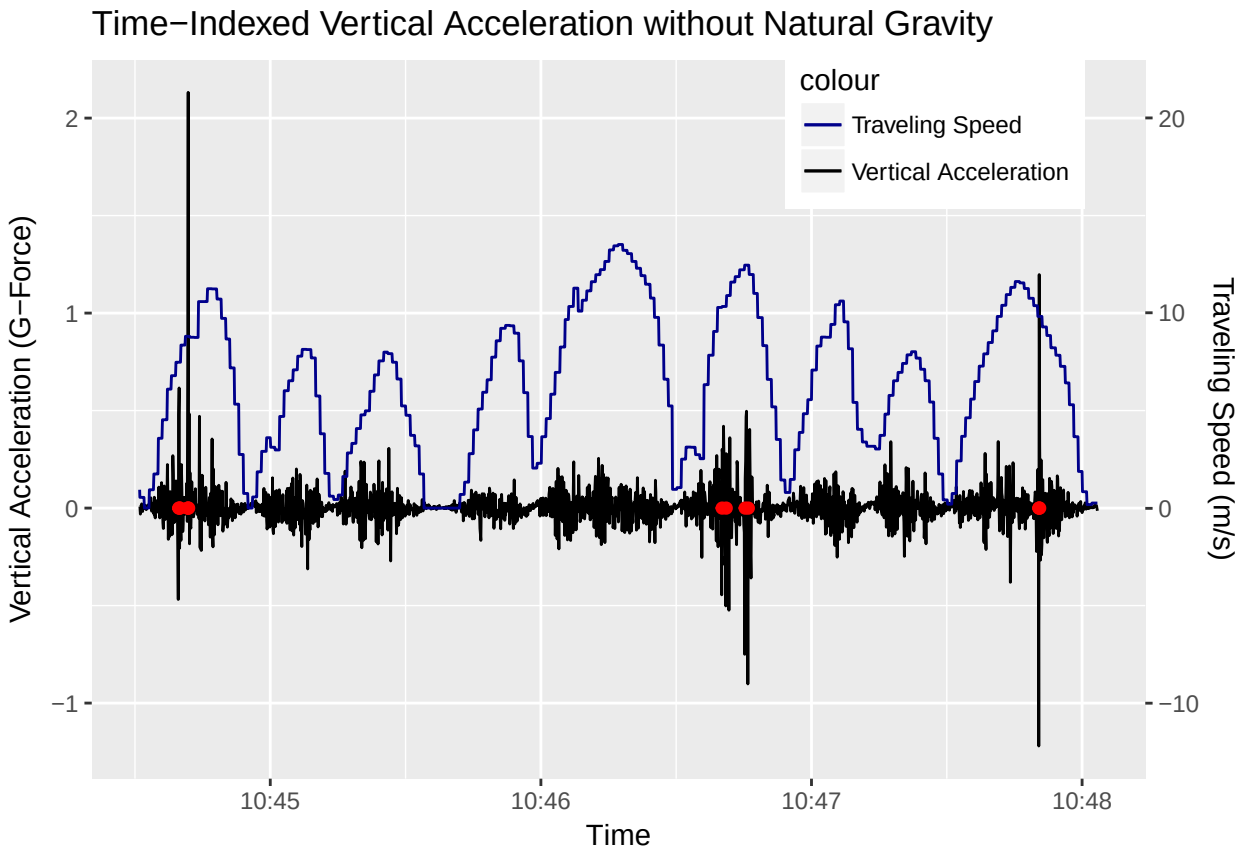


Note: Speed bumps are labeled as *red* points on the graph.

Epoch 2: Display of Vertical Acceleration without Natural Gravity

vert_accel: vertical acceleration (up and down) of the vehicle without natural gravity

```
# remove natural gravity in vertical acceleration
LA10.valid$vert_accel = LA10.valid$vert_accel_G + 1
```

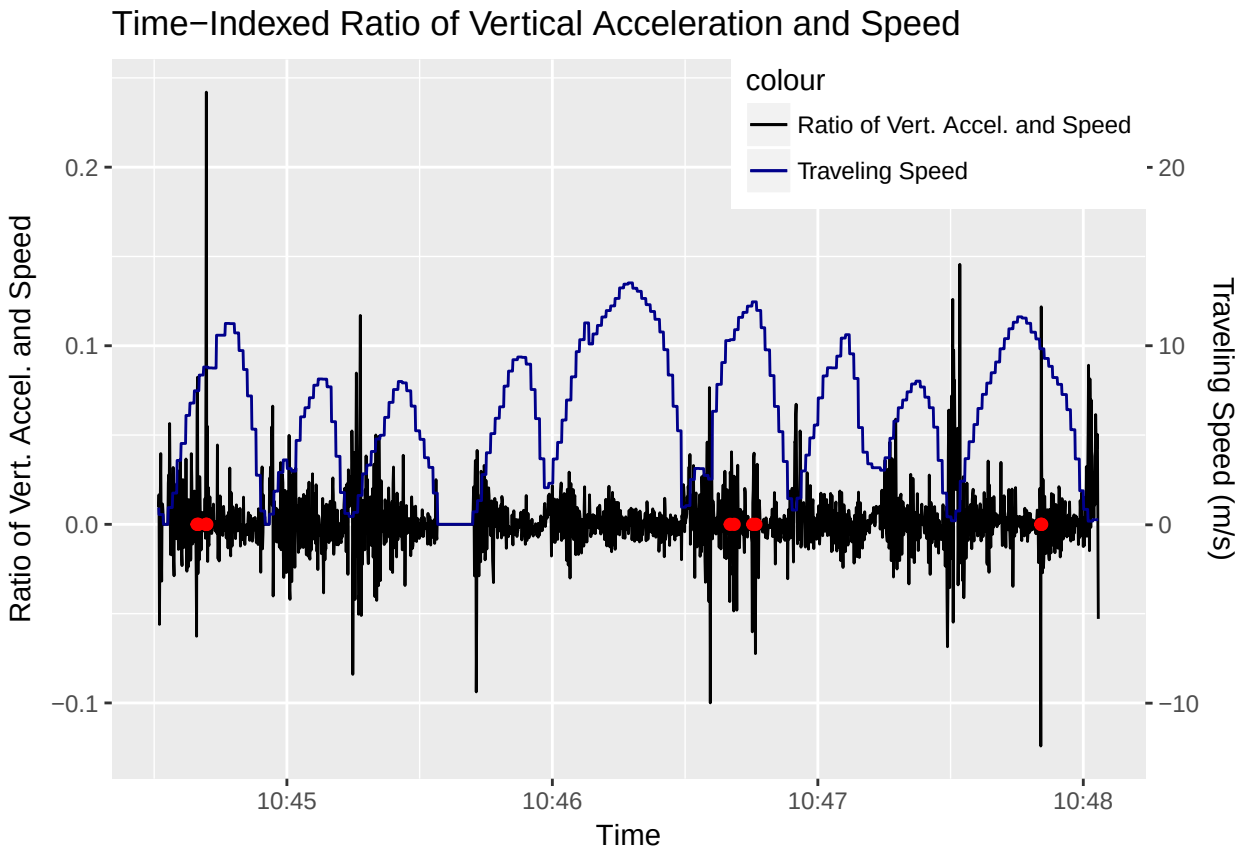


Note: Speed bumps are labeled as *red* points on the graph.

Epoch 3: Display of Ratio between Vertical Acceleration (w/o Natural G) and Speed

vert_accel_ratio_speed: the ratio between vertical acceleration (without natural gravity) and traveling speed

```
# calculate ratio between vertical acceleration and speed
for (i in 1:nrow(LA10.valid)) {
    if (LA10.valid$Speed[i] == 0) {
        LA10.valid$vert_accel_ratio_speed[i] = 0
    }
    else {
        LA10.valid$vert_accel_ratio_speed[i] = LA10.valid$vert_accel[i] / LA10.valid$Speed[i]
    }
}
```

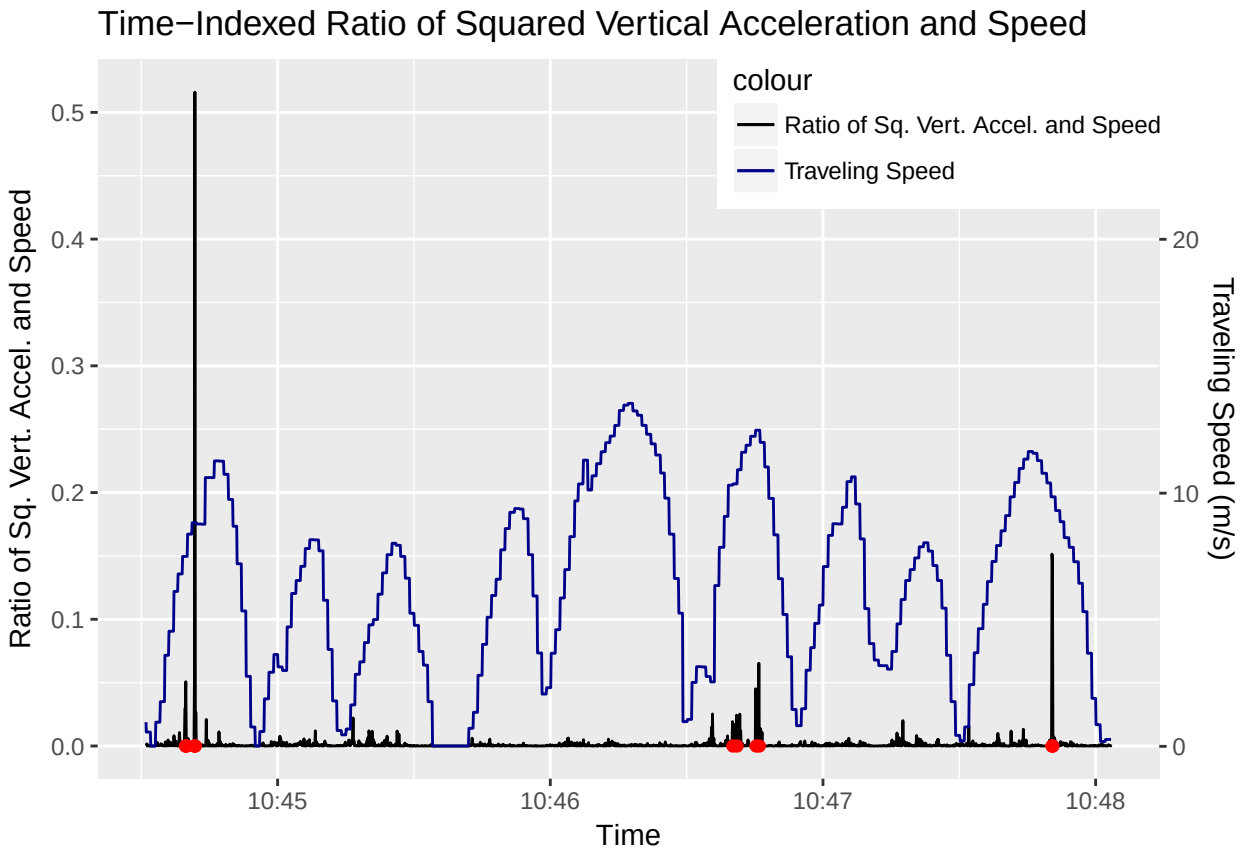


Note: Speed bumps are labeled as *red* points on the graph.

Epoch 4: Display of Ratio between Vertical Acceleration (w/o Natural G) and Speed

sq_vert_accel_ratio_speed: the ratio between squared vertical acceleration (without natual gravity) and traveling speed

```
# calculate ratio between vertical acceleration and speed
for (i in 1:nrow(LA10.valid)) {
    if (LA10.valid$Speed[i] == 0) {
        LA10.valid$sq_vert_accel_ratio_speed[i] = 0
    }
    else {
        LA10.valid$sq_vert_accel_ratio_speed[i] = (LA10.valid$vert_accel[i] * LA10.valid$vert_accel[i])
    }
}
```

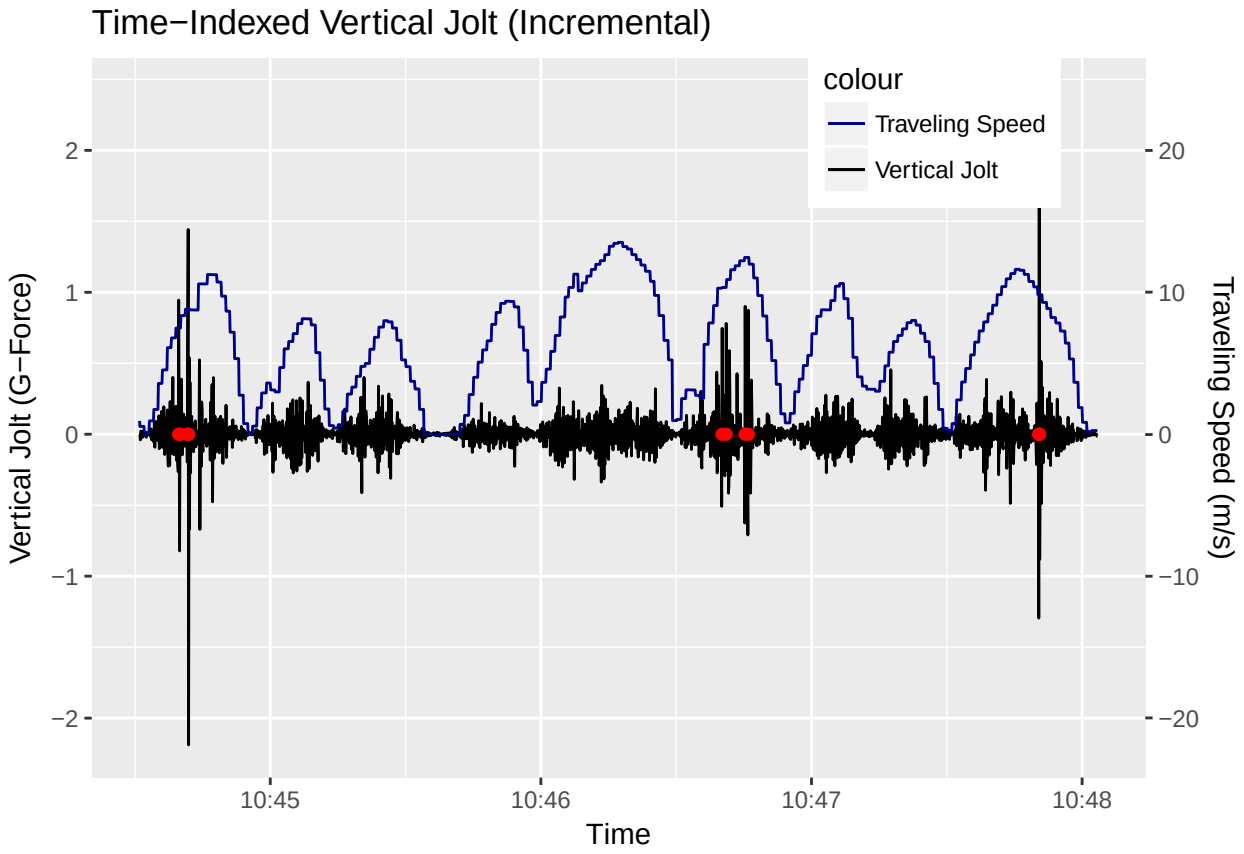


Note: Speed bumps are labeled as *red* points on the graph.

Epoch 5: Display of Vertical Jolt (Incremental Change of Vert. Accel.)

vert_jolt: vertical jolt of the vehicle (incremental change of vertical acceleration)

```
# calculate vertical jolt
for (i in 1:nrow(LA10.valid)) {
    if (i == 1) {
        LA10.valid$vert_jolt[i] = 0
    }
    else {
        LA10.valid$vert_jolt[i] = LA10.valid$vert_accel[i] - LA10.valid$vert_accel[i - 1]
    }
}
```

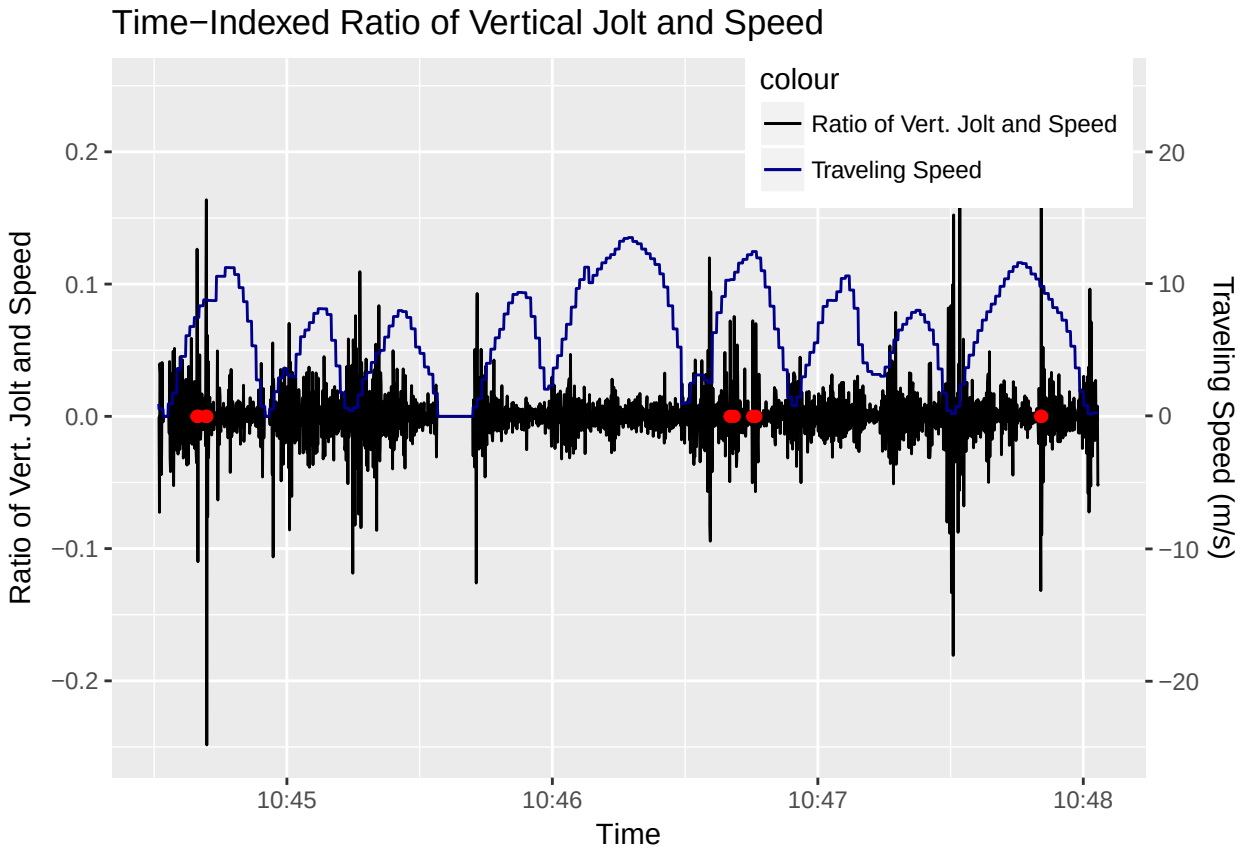


Note: Speed bumps are labeled as *red* points on the graph.

Epoch 6: Display of Ratio between Vertical Jolt and Speed

vert_jolt_ratio_speed: the ratio between vertical jolt and traveling speed

```
# calculate ratio between vertical jolt and speed
for (i in 1:nrow(LA10.valid)) {
    if (LA10.valid$Speed[i] == 0) {
        LA10.valid$vert_jolt_ratio_speed[i] = 0
    }
    else {
        LA10.valid$vert_jolt_ratio_speed[i] = LA10.valid$vert_jolt[i] / LA10.valid$Speed[i]
    }
}
```

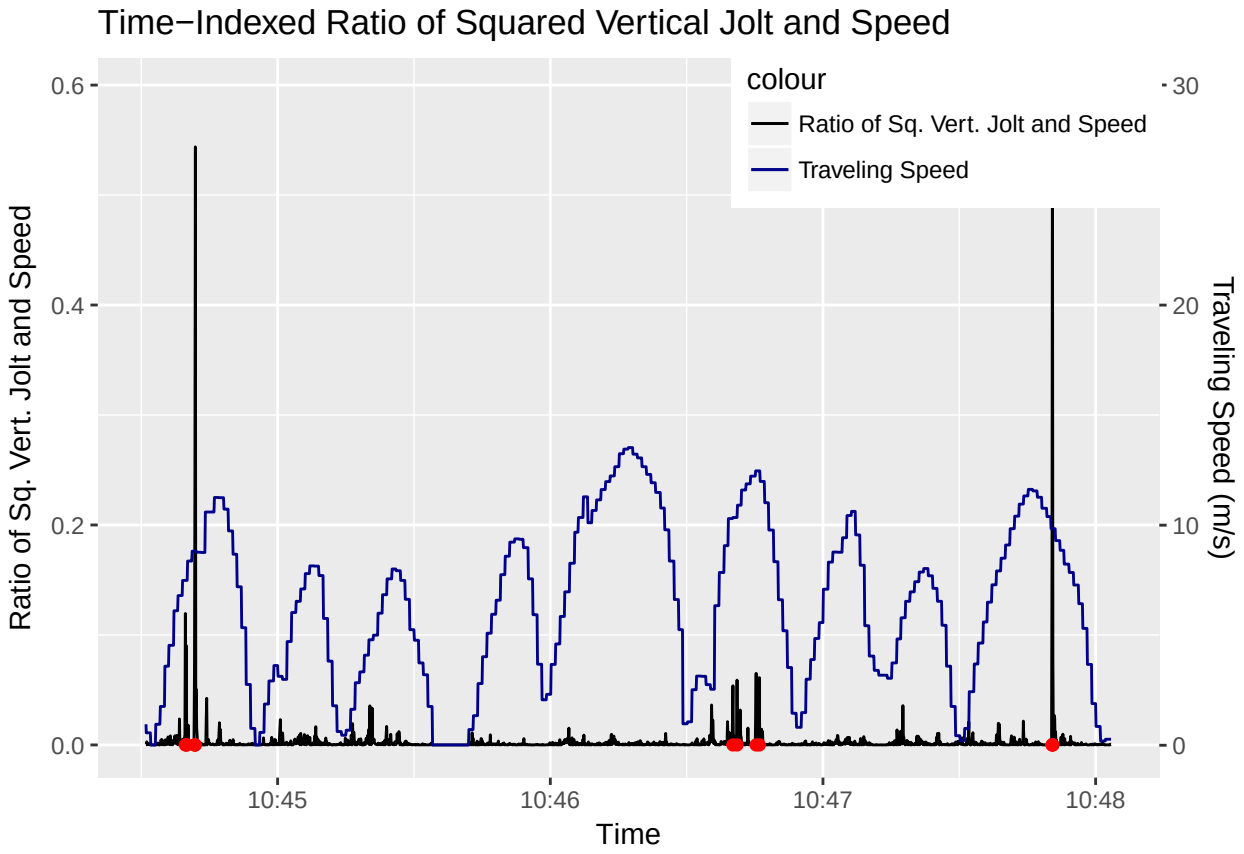


Note: Speed bumps are labeled as *red* points on the graph.

Epoch 7: Display of Ratio between Squared Vertical Jolt and Speed

sq_vert_jolt_ratio_speed: the ratio between squared vertical jolt and traveling speed

```
# calculate ratio between vertical jolt and speed
for (i in 1:nrow(LA10.valid)) {
    if (LA10.valid$Speed[i] == 0) {
        LA10.valid$sq_vert_jolt_ratio_speed[i] = 0
    }
    else {
        LA10.valid$sq_vert_jolt_ratio_speed[i] = (LA10.valid$vert_jolt[i] * LA10.valid$vert_jolt[i]) / I
    }
}
```



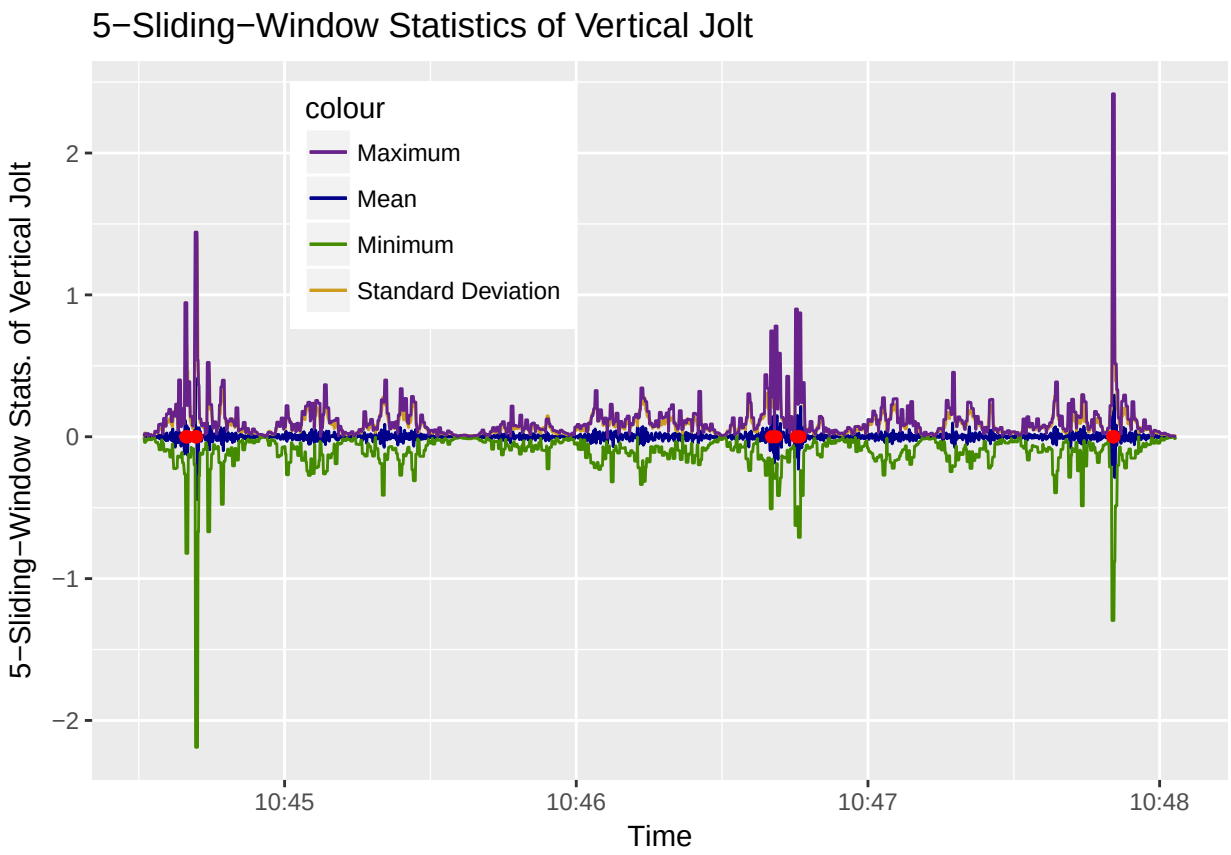
Note: Speed bumps are labeled as *red* points on the graph.

Epoch 8: Display of Sliding-Window Statistics of Vertical Jolt

vert_jolt_mean: 5-sliding-window mean of vertical jolt
vert_jolt_sd: 5-sliding-window standard deviation of vertical jolt
vert_jolt_min: 5-sliding-window minimum of vertical jolt
vert_jolt_max: 5-sliding-window maximum of vertical jolt

```
# calculate 5-sliding-window mean of vertical jolt
for (i in 3:nrow(LA10.valid)-2) {
    LA10.valid$vert_jolt_mean[i] = mean(c(LA10.valid$vert_jolt[i-2],
                                          LA10.valid$vert_jolt[i-1],
                                          LA10.valid$vert_jolt[i],
                                          LA10.valid$vert_jolt[i+1],
                                          LA10.valid$vert_jolt[i+2]))
}
LA10.valid$vert_jolt_mean[1] = 0
LA10.valid$vert_jolt_mean[2] = 0
LA10.valid$vert_jolt_mean[nrow(LA10.valid) - 1] = 0
LA10.valid$vert_jolt_mean[nrow(LA10.valid)] = 0
# calculate 5-sliding-window standard deviation of vertical jolt
for (i in 3:nrow(LA10.valid)-2) {
    LA10.valid$vert_jolt_sd[i] = sd(c(LA10.valid$vert_jolt[i-2],
                                      LA10.valid$vert_jolt[i-1],
                                      LA10.valid$vert_jolt[i],
                                      LA10.valid$vert_jolt[i+1],
                                      LA10.valid$vert_jolt[i+2]))
}
LA10.valid$vert_jolt_sd[1] = 0
LA10.valid$vert_jolt_sd[2] = 0
LA10.valid$vert_jolt_sd[nrow(LA10.valid) - 1] = 0
LA10.valid$vert_jolt_sd[nrow(LA10.valid)] = 0
# calculate 5-sliding-window minimum of vertical jolt
for (i in 3:nrow(LA10.valid)-2) {
    LA10.valid$vert_jolt_min[i] = min(c(LA10.valid$vert_jolt[i-2],
                                        LA10.valid$vert_jolt[i-1],
                                        LA10.valid$vert_jolt[i],
                                        LA10.valid$vert_jolt[i+1],
                                        LA10.valid$vert_jolt[i+2]))
}
LA10.valid$vert_jolt_min[1] = 0
LA10.valid$vert_jolt_min[2] = 0
LA10.valid$vert_jolt_min[nrow(LA10.valid) - 1] = 0
LA10.valid$vert_jolt_min[nrow(LA10.valid)] = 0
# calculate 5-sliding-window maxium of vertical jolt
for (i in 3:nrow(LA10.valid)-2) {
    LA10.valid$vert_jolt_max[i] = max(c(LA10.valid$vert_jolt[i-2],
                                        LA10.valid$vert_jolt[i-1],
                                        LA10.valid$vert_jolt[i],
                                        LA10.valid$vert_jolt[i+1],
                                        LA10.valid$vert_jolt[i+2]))
}
LA10.valid$vert_jolt_max[1] = 0
LA10.valid$vert_jolt_max[2] = 0
```

```
LA10.valid$vert_jolt_max[nrow(LA10.valid) - 1] = 0
LA10.valid$vert_jolt_max[nrow(LA10.valid)] = 0
```



Note: Speed bumps are labeled as *red* points on the graph.

Data Writing

Date: the time stamp of a data point
Speed: traveling speed of the vehicle
forw_accel: forward acceleration (front and back) of the vehicle
hori_accel: horizontal acceleration (left and right) of the vehicle
vert_accel_G: vertical acceleration (up and down) of the vehicle with natural gravity (-1G)
speedbump: whether this data point is a speedbump
vert_accel: vertical acceleration (up and down) of the vehicle without natural gravity
vert_accel_ratio_speed: the ratio between vertical acceleration (without natural gravity) and traveling speed
sq_vert_accel_ratio_speed: the ratio between squared vertical acceleration (without natual gravity) and traveling speed
vert_jolt: vertical jolt of the vehicle (incremental change of vertical acceleration)
vert_jolt_ratio_speed: the ratio between vertical jolt and traveling speed
sq_vert_jolt_ratio_speed: the ratio between squared vertical jolt and traveling speed
vert_jolt_mean: 5-sliding-window mean of vertical jolt
vert_jolt_sd: 5-sliding-window standard deviation of vertical jolt
vert_jolt_min: 5-sliding-window minimum of vertical jolt
vert_jolt_max: 5-sliding-window maximum of vertical jolt

```
jpeg("./Epoch1_LA10.vert_accel_G.plot.jpeg", width = 1200, height = 750, units = "px")
print(LA10.vert_accel_G.plot)
dev.off()
```

```
## pdf
##   2
```

```
jpeg("./Epoch2_LA10.vert_accel.plot.jpeg", width = 1200, height = 750, units = "px")
print(LA10.vert_accel.plot)
dev.off()
```

```
## pdf
##   2
```

```
jpeg("./Epoch3_LA10.vert_accel_ratio_speed.plot.jpeg", width = 1200, height = 750, units = "px")
print(LA10.vert_accel_ratio_speed.plot)
dev.off()
```

```
## pdf
##   2
```

```
jpeg("./Epoch4_LA10.sq_vert_accel_ratio_speed.plot.jpeg", width = 1200, height = 750, units = "px")
print(LA10.sq_vert_accel_ratio_speed.plot)
dev.off()
```

```
## pdf
##   2
```

```
jpeg("./Epoch5_LA10.vert_jolt.plot.jpeg", width = 1200, height = 750, units = "px")
print(LA10.vert_jolt.plot)
dev.off()
```

```
## pdf
##   2
```

```
jpeg("./Epoch6_LA10.vert_jolt_ratio_speed.plot.jpeg", width = 1200, height = 750, units = "px")
print(LA10.vert_jolt_ratio_speed.plot)
dev.off()
```

```
## pdf
##   2
```

```
jpeg("./Epoch7_LA10.sq_vert_jolt_ratio_speed.plot.jpeg", width = 1200, height = 750, units = "px")
print(LA10.sq_vert_jolt_ratio_speed.plot)
dev.off()
```

```
## pdf
##   2
```

```
jpeg("./Epoch8_LA10.vert_jolt_5.plot.jpeg", width = 1200, height = 750, units = "px")
print(LA10.vert_jolt_5.plot)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(LA10.valid, "./los_angeles_10_labeled.csv")
```