

Formule 1 | Grand Prix de Monaco 2021 | Fernando Alonso

Baptiste Gorteau & Théo Cavenne

Description des données

Pour ce projet, nous avons décidé d'étudier la vitesse de Fernando Alonso lors du Grand Prix de Formule 1 de Monaco en 2021.

Les jeux de données que nous allons utiliser sont les suivant :

- “laps.csv” : Récapitulatif avec différentes informations pour chaque tour effectué par Fernando Alonso (temps au tour, numéro du tour, type de pneus utilisées, etc.). Les seules variables que nous allons utiliser dans ce jeu de données sont “LapNumber” (numéro du tour), “PitOutTime” et “PitInTime” (moments où Alonso est entré ou sorti des stands). On nommera ce jeu de données “laps_alonso” :

```
##      LapNumber      PitOutTime      PitInTime
## 1           1 0 days 00:00:04.827000
## 45          45              0 days 01:31:34.693000
## 46          46 0 days 01:31:58.267000
```

- “lap_i.0.csv” (pour i allant de 1 à 77) : On aura ici 77 jeux de données différents représentant chacun un tour du GP 2021 de Monaco effectué par Fernando Alonso. Pour chacun de ces jeux de données (tours), on aura plusieurs couples Distance/Vitesse qui représentent la vitesse mesurée à une distance précise de la ligne de départ pour un tour. On va donc concaténer ces 77 jeux de données afin de n'avoir qu'un seul dataframe (24861 lignes et 3 colonnes) composée des variables “Distance” (distance en mètres par rapport à la ligne de départ du points de mesure de la vitesse), “Speed” (vitesse en km/h mesurée) et “lap” (numéro du tour (1 à 77)). On nommera ce jeu de données “ALO” :

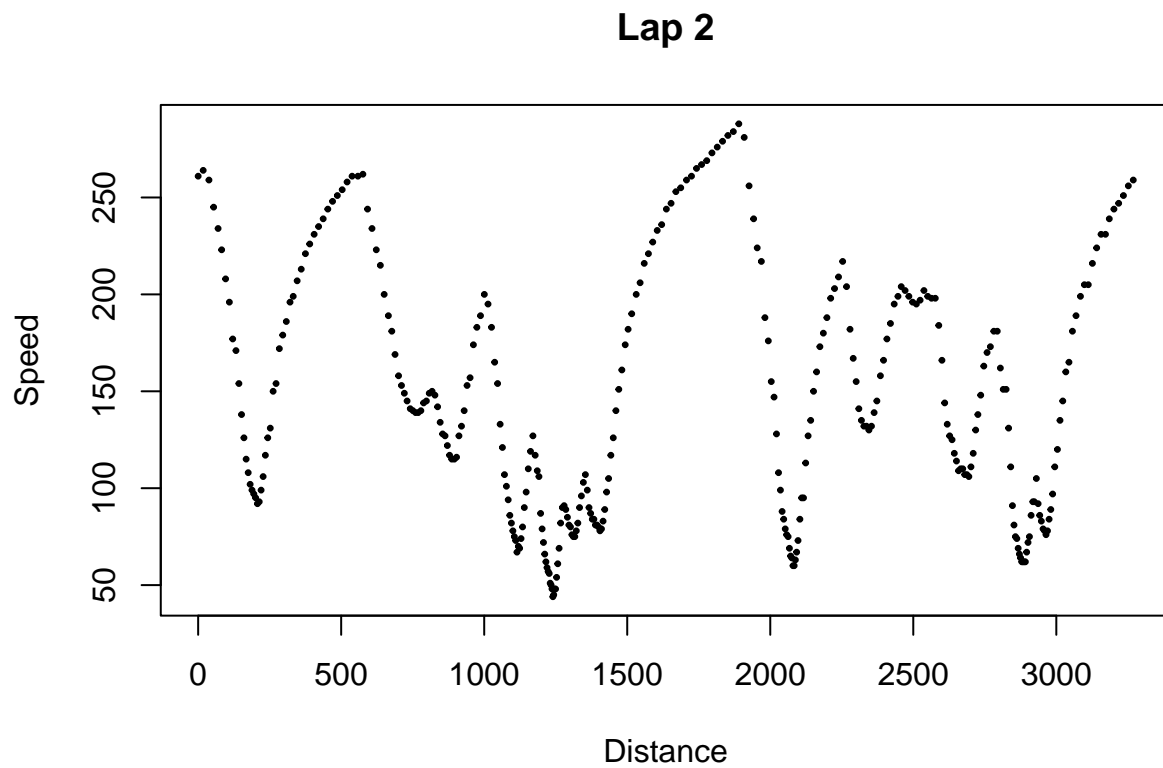
```
##      Distance Speed lap
## 1 0.000000      0    1
## 2 0.000000      0    1
## 3 0.000000      0    1
## 4 1.000000     15    1
## 5 2.466667     22    1
## 6 4.466667     30    1
```

Pour faciliter notre étude, nous avons décidé d'enlever le premier tour et les tours de changements de pneus de notre jeu de données car on estime qu'ils ne sont pas représentatifs d'un tour “normal” d'un Grand Prix.

Indices des tours concernés :

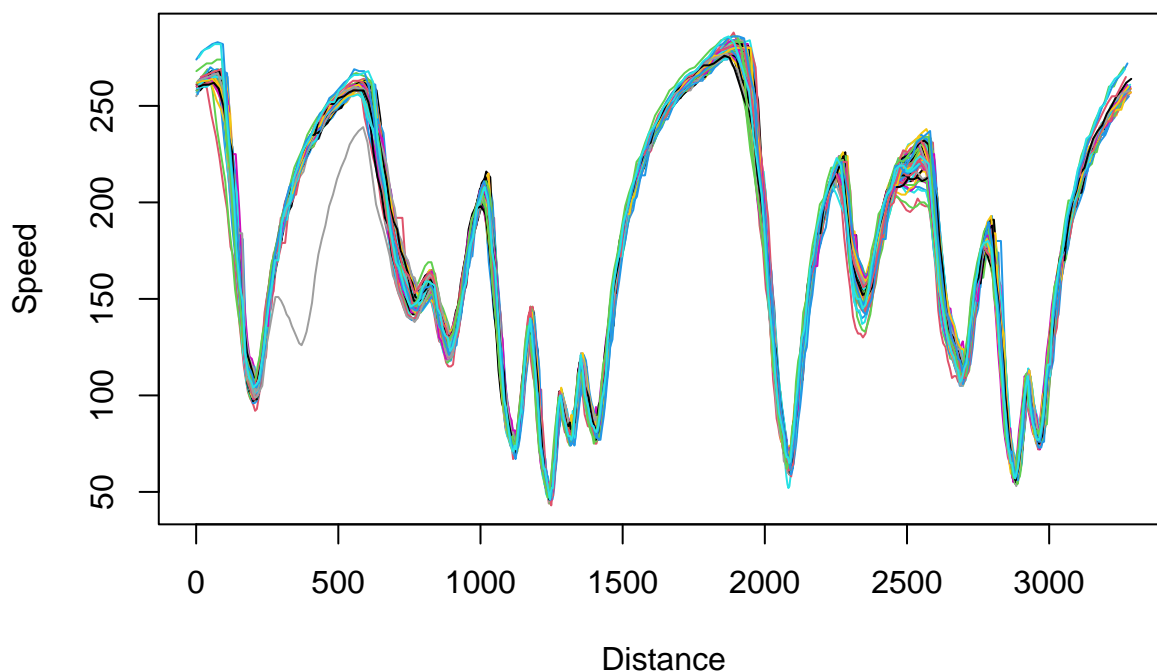
```
## [1] 1 46 45
```

Visualisation de la courbe de la vitesse (en fonction de la distance à la ligne de départ) du 2ème tour :



Visualisation des courbes de vitesse pour tous les tours :

Courbes observées



Les représentations de la vitesse pour chaque tour semblent être régulières et non-périodiques. Nous avons donc décidé de ne pas étudier les bases de Fourier ainsi que les bases d'ondelettes. Nous allons, pour la suite, nous concentrer sur les bases de splines cubiques.

Choix des noeuds de la base de splines cubique

Première base

Sur la représentation des courbes observées ci-dessus, on remarque qu'il y a 24 changements "brutaux" (pics) entre accélération et decélération. On a donc décidé de choisir nos noeuds "à la main". Les 24 noeuds choisis représentent les distances des entrées et des sorties de ces virages (par rapport à la ligne de départ).

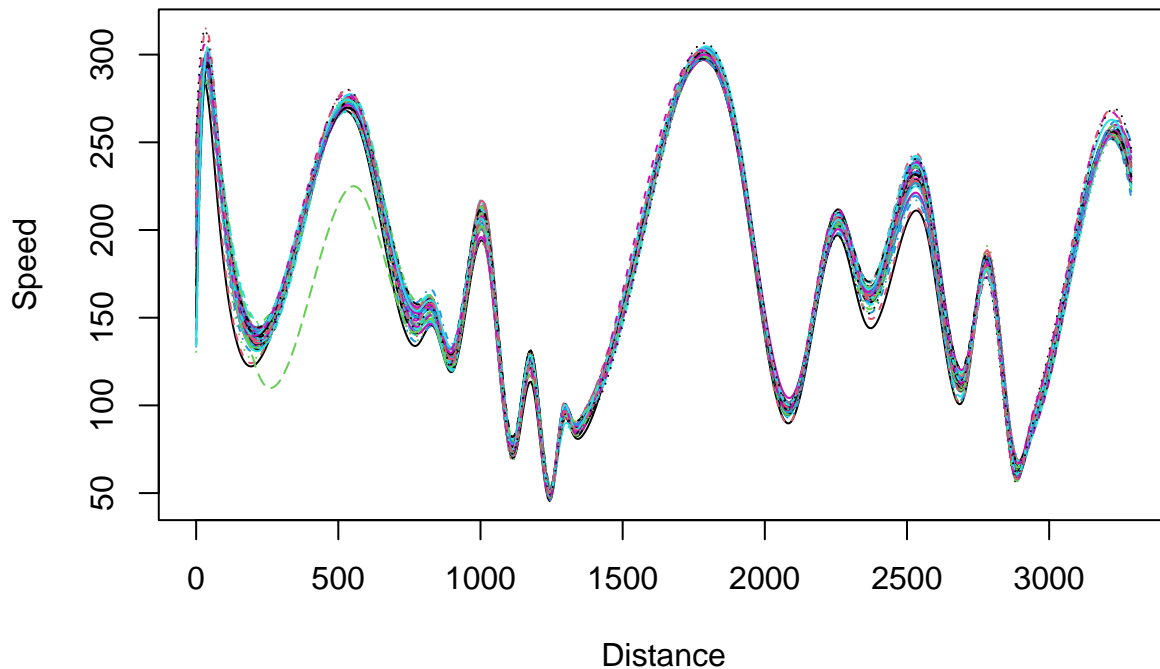
- Projection des courbes individuelles sur la base de splines choisie :

```
##          [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]      [,8]
## [1,] 249.4464 329.2739 133.0935 85.56907 402.9495 96.49027 161.4706 72.05380
## [2,] 241.8686 346.3732 134.5251 88.41474 399.1156 109.64074 159.7720 83.59400
## [3,] 148.9952 380.4808 144.0943 82.03570 400.7732 100.91267 162.2687 75.64593
## [4,] 212.3866 360.7799 151.0482 82.75027 400.3937 103.49681 158.9545 77.24844
## [5,] 230.8997 349.6562 158.6731 82.33900 398.6708 108.98015 160.5046 84.30591
## [6,] 192.7531 361.2370 158.4142 81.80166 404.8370 112.20452 158.6068 81.74960
##          [,9]      [,10]      [,11]      [,12]      [,13]      [,14]      [,15]
## [1,] 291.9376 -0.5914849 169.3460 -2.18201696 109.9191 78.42846 82.15088
## [2,] 290.2612 -0.9991253 174.2381 0.02185931 112.3124 80.71636 88.16135
## [3,] 289.1854 -5.6396762 177.9938 -1.85449445 110.9173 77.90655 89.28481
## [4,] 293.6899 -4.5641840 185.4453 -8.82971619 119.3539 77.70835 92.27863
## [5,] 290.0611 -0.9729180 178.8423 -5.18072192 121.0063 77.31326 89.50983
```

```
## [6,] 291.0734 -1.0318486 181.0933 -5.08732413 117.7785 76.50881 91.64160
##      [,16]      [,17]      [,18]      [,19]      [,20]      [,21]      [,22]      [,23]
## [1,] 159.8242 475.8930 -18.33004 261.2234 72.10967 323.3761 16.82985 263.6750
## [2,] 156.9318 477.7771 -20.96782 268.9931 77.37146 322.8672 21.01838 253.0340
## [3,] 149.3189 481.8406 -10.32344 256.6620 86.16271 331.8148 15.13709 270.0005
## [4,] 149.3264 490.2594 -15.29737 258.2452 86.05361 326.3855 21.05396 257.4330
## [5,] 163.2045 471.4013 -13.94253 269.6682 87.99255 331.8714 22.41257 260.1666
## [6,] 147.2648 483.2113 -10.84651 260.5365 95.18182 327.1103 31.69031 251.7197
##      [,24]      [,25]      [,26]      [,27]      [,28]
## [1,] 24.91328 82.83943 141.5979 311.3169 233.3012
## [2,] 31.44112 80.10483 132.0638 304.7344 241.4852
## [3,] 41.10768 80.11761 130.1162 302.6818 235.8910
## [4,] 39.50774 77.49367 138.5440 307.0326 233.8692
## [5,] 33.08187 79.65240 146.1441 303.3999 236.2998
## [6,] 25.06016 84.14245 144.8551 308.6503 232.1546
```

- Visualisation des courbes lissées :

Courbes lissées



```
## [1] "done"
```

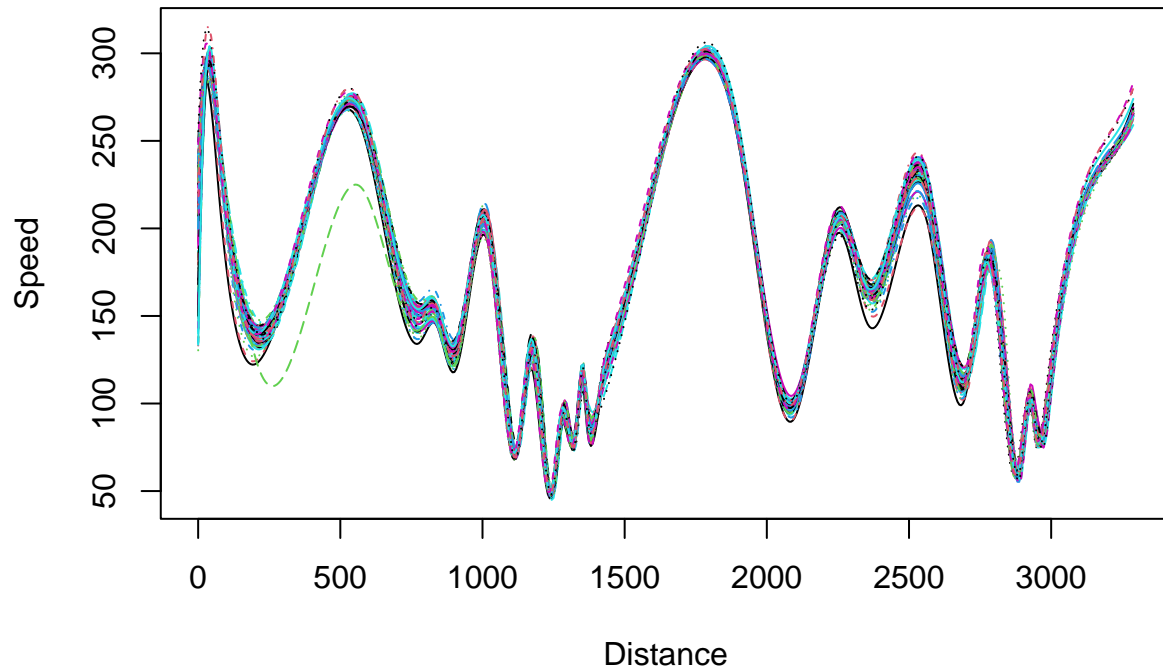
On observe qu'il manque 2 changements "brutaux" entre accélération et decélération" entre 1100 et 1450 ainsi que 2 entre 2275 et 3000.

Deuxième base

On va donc réutiliser la même base en ajoutant 5 noeuds supplémentaires dans chacun de ces intervalles (on ajoutera donc 10 noeuds au total). On aura dans ce cas ci une base de splines d'ordre 4 avec 34 noeuds intérieurs et donc 38 fonctions de bases.

- Visualisation des courbes lissées :

Courbes lissées



```
## [1] "done"
```

Ces représentations nous semblent très satisfaisantes. Nous allons donc retenir la base de splines créée ci-dessus pour la suite.

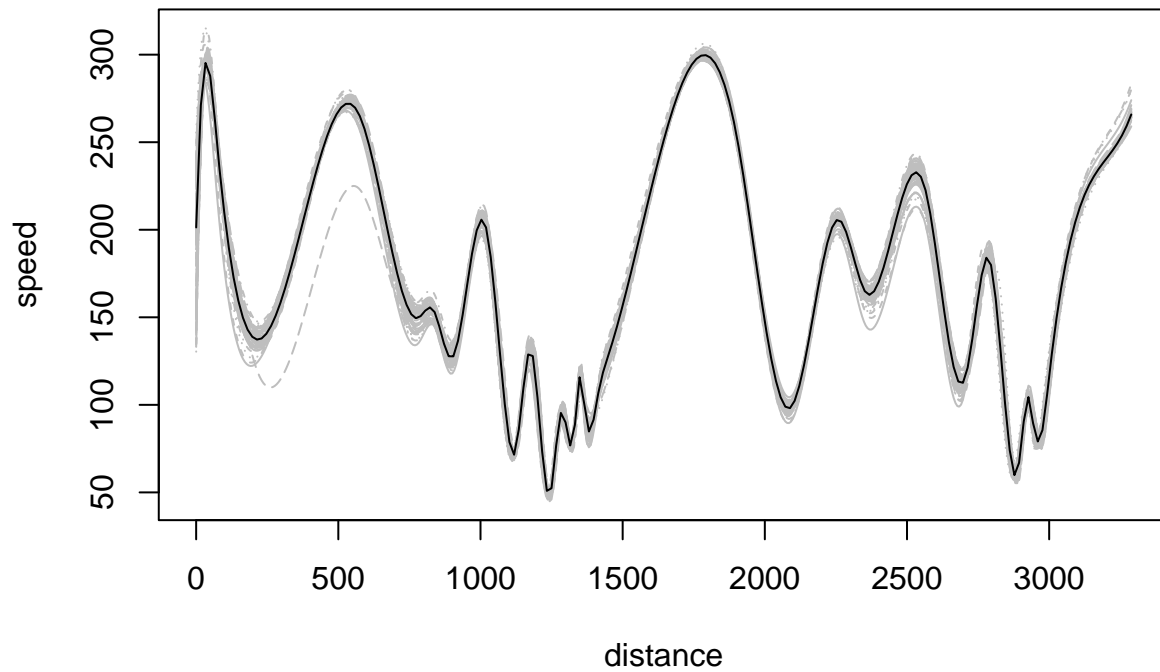
Statistiques exploratoires

Étude de la moyenne et de la variance

Moyenne

```
## [1] "done"
```

Moyenne des tours d'Alonso

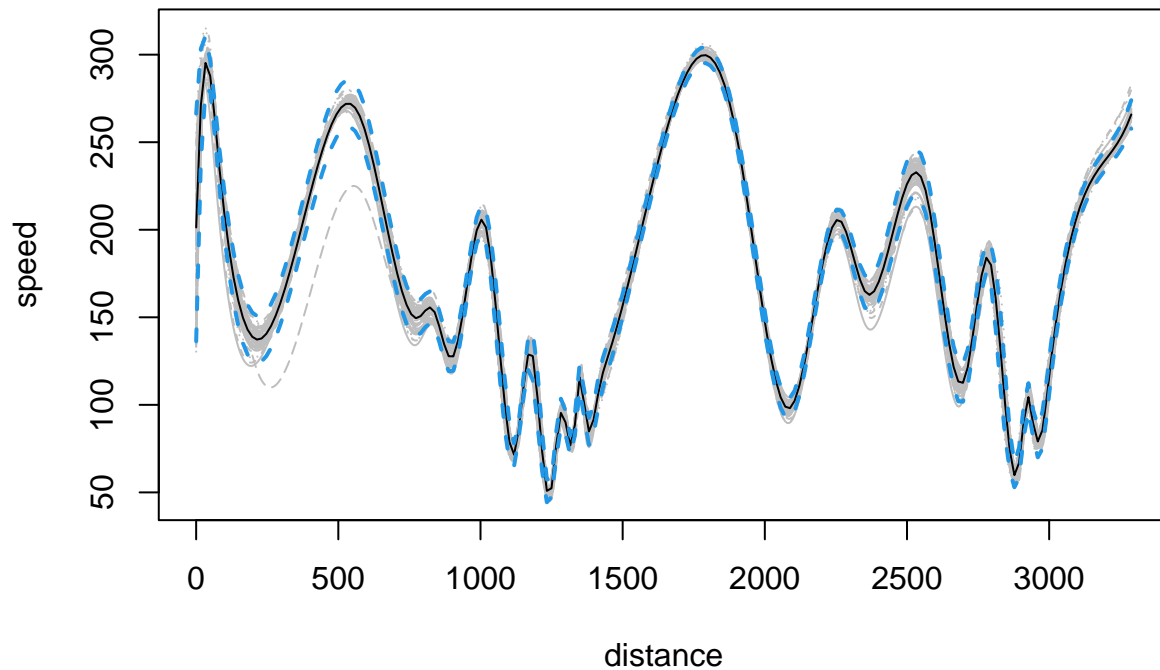


Comme nous pouvions nous y attendre, la moyenne est proche de toutes les données. Cette moyenne représente donc très bien un tour de Fernando Alonso lors du Grand Prix de Monaco. À présent, nous allons regarder la variance des tours.

Variance

```
## [1] "done"
```

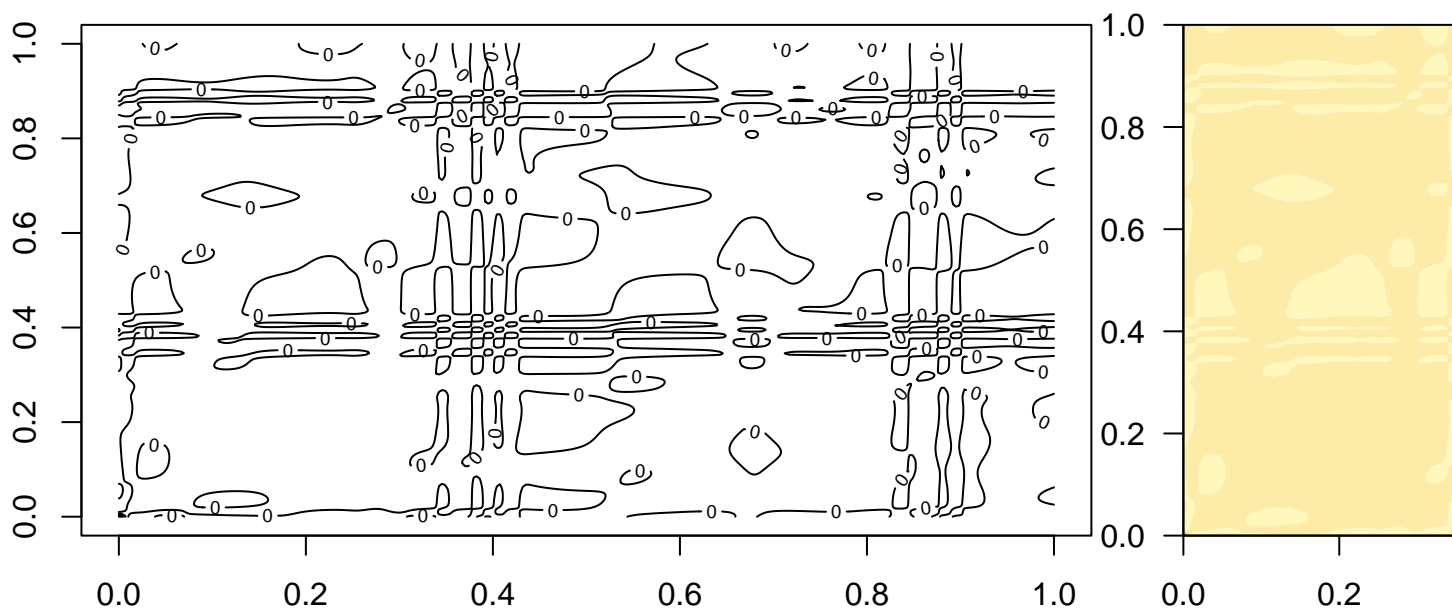
Moyenne des tours d'Alonso



Nous remarquons sur ce graphique que quasiment toutes les parties des tours se trouvent dans l'intervalle de la variance. On peut observer une variance plus forte lors des moments de freinage et de réaccélération. Au contraire, lors des moments d'accélération, la vitesse varie très peu selon les tours. Il est intéressant de noter que la variance au freinage est plus élevée lorsque le freinage est fort (comparaison entre le freinage à 1700m et celui à 2500m).

Covariance et corrélation

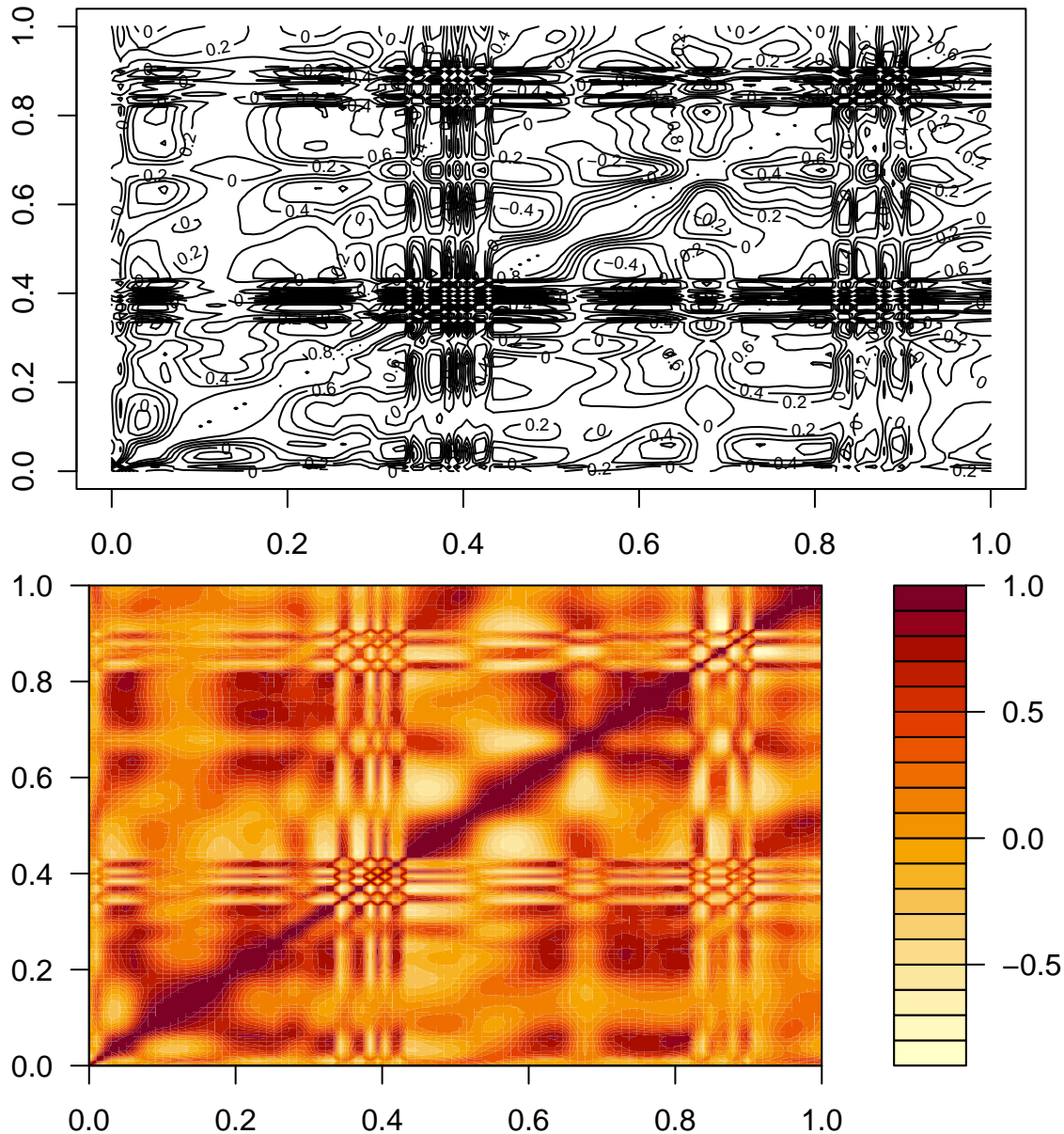
Covariance



Nous rencontrons un problème que nous n'arrivons pas à résoudre. Au niveau des valeurs très proche de zéro, la covariance est extrêmement élevée. Nous décidons donc de restreindre la représentation de la covariance aux valeurs qui ne sont pas proches de ce pic abérant.

Sur cette représentation on remarque des covariances plutôt faibles. Cependant, on remarque un pic vers la fin du tour (mesure 278). On peut donc penser qu'il y a une forte variation des points entre eux sur la fin du circuit. Nous pouvons penser que ce point correspond au freinage au bout de 2.5km. Les variations au niveau de ce point sont observables sur la représentation de la variance plus haut.

Corrélation



Grâce à ces graphiques, on peut clairement voir qu'il y a une forte corrélation entre des points qui se suivent. Un point x et $x+1$ vont être très corrélés entre eux. Ce qui est logique à la vue de nos données.

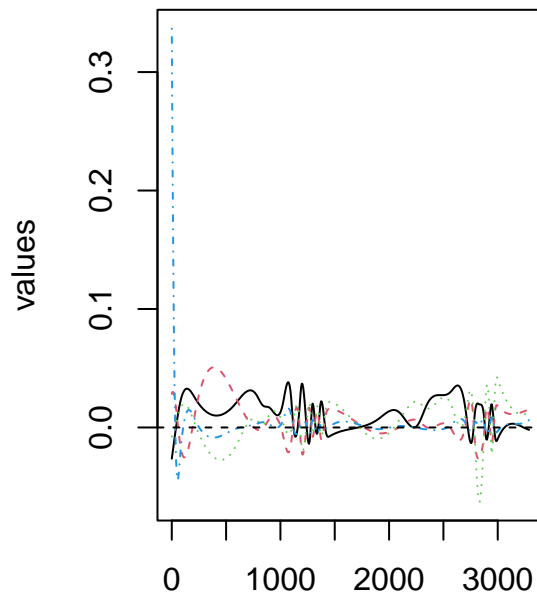
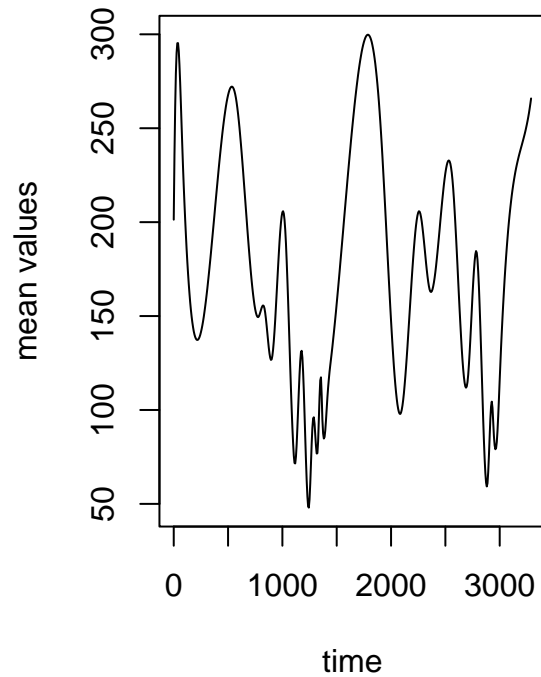
ACP fonctionnelle

```
## [1] 0.3705133 0.1664614 0.1396466 0.1240051
```

```
## [1] 0.3705133 0.5369747 0.6766214 0.8006264
```

Ici, les deux premières composantes principales expliquent 64% des données et les quatre premières en expliquent 82%.

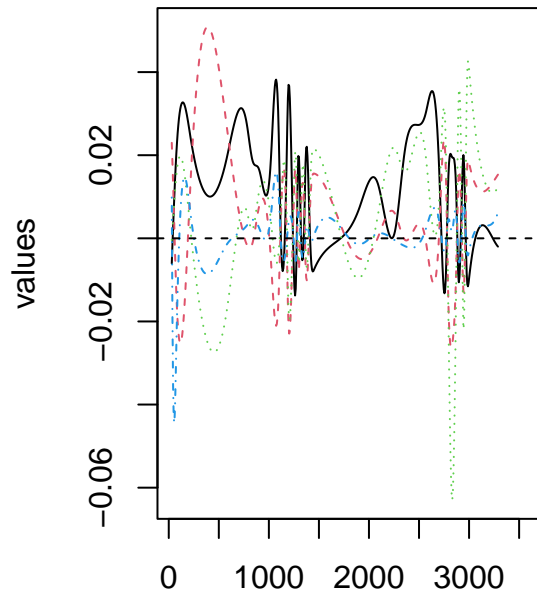
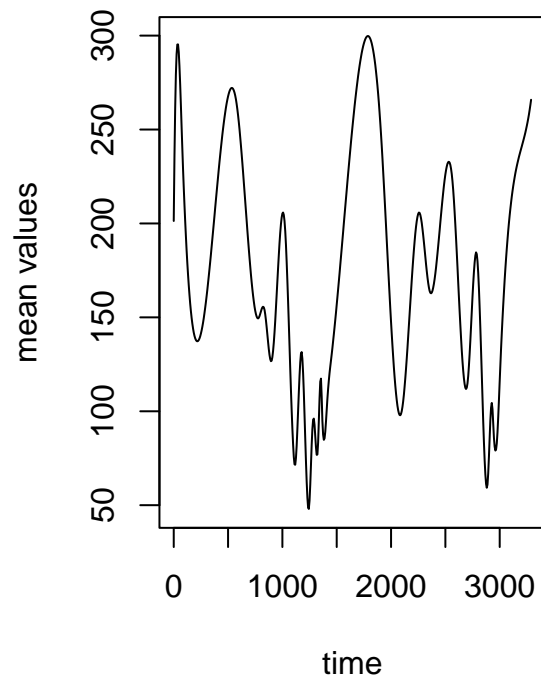
```
## [1] "done"
```



```
## [1] "done"
```

On rencontre un problème avec la troisième composante qui est très élevé en 0, on décide de réduire

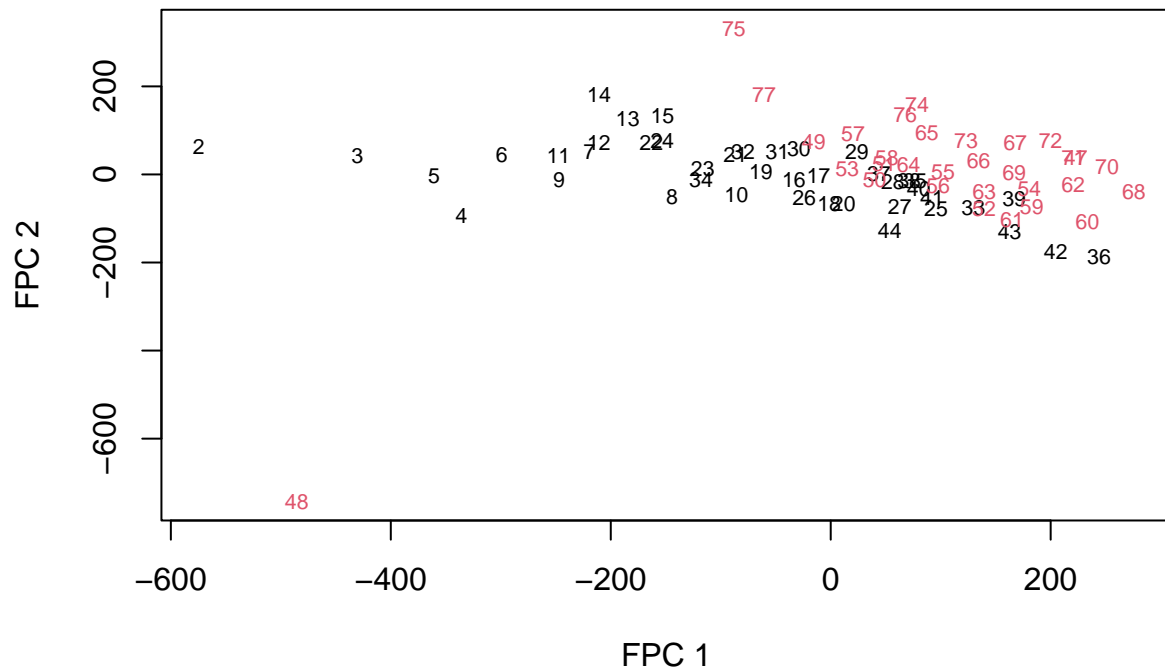
```
## [1] "done"
```



```
## [1] "done"
```

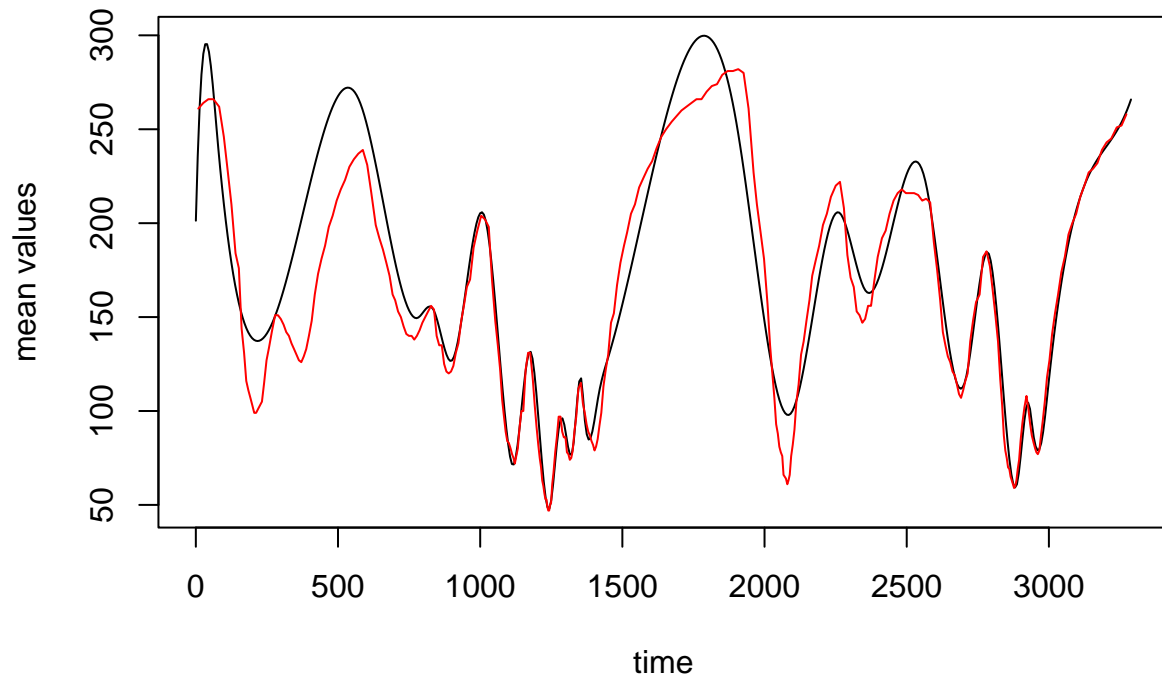
Nous pouvons observer sur le graphique de droite les représentations des composantes principales de l'ACP réalisée. La courbe noir et la courbe rouge en pointillé sont celles qui représentent 64% des données. La courbe noir sépare surtout les données en fonction du premier kilomètre et du secteur aux alentours de 2,5km.

##	[,1]	[,2]	[,3]	[,4]
## [1,]	-574.5493	64.457571	-217.81712	100.235748
## [2,]	-430.4633	42.499948	-254.24174	91.328248
## [3,]	-335.6875	-93.061531	-281.90046	-173.696881
## [4,]	-360.7013	-2.790562	-208.50670	2.876585
## [5,]	-299.0054	44.124505	-84.32063	60.633913
## [6,]	-219.3278	52.304805	-95.57406	-53.531394



Ce graphique permet de visualiser dans un premier temps que l'axe 2 sépare le tour 48 des autres tours. Nous allons donc regarder quel était ce tour.

```
## [1] "done"
```



En effet, on remarque que sur ce tour, Fernando Alonso a freiné anormalement dès le début du tour (400m). De plus, le pic du deuxième axe principal correspond à cette distance d'environ 400m.

C'est ce freinage qui fait que ce tour est différent des autres.

Si on revient à la projection des tours sur les axes, on remarque que l'axe 1 sépare plutôt bien les tours effectués en pneu medium (couleur noir) de ceux effectués en pneu soft.

Lorsqu'on reprend la représentation de la variance et l'analyse qui en avait été faite, on peut dire que le type

de pneu va avoir une influence sur la vitesse de freinage et sur la vitesse de réaccélération.

De ce que nous avons pu observer, les moments de freinage à 250m mais aussi à 800m et 2.5km sont sûrement les points du circuit où Fernando Alonso peut gagner ou perdre du temps sur son tour. le type de pneu va également jouer sur son freinage à ces instants.