

TP - Analyse subjective-objective de signaux audio

Yassine Jamoud, Samy Haffoudhi

3 février 2022

Introduction

Le « bruit de toqué » est obtenu lorsqu'un client « sonne » la planche de bord d'une voiture en frappant dessus avec ses doigts. Il tire de cette information une évaluation subjective de la qualité perçue. L'objectif du TP est de comprendre ce qui fonde, du point de vue du signal, les différences perceptives entre différentes planches, et d'expliquer les préférences utilisateur. Le but ultime de l'étude est de savoir comment reconcevoir la planche de bord (matériaux, montage) pour augmenter la préférence utilisateur.

Nous disposons de 16 sons enregistrés dans des conditions identiques sur différents véhicules. Ces sons ont été évalués par un panel d'experts sensoriels selon 5 descripteurs :

- hauteur
- détonnant
- attaquant
- éloignement
- intensité

1 Écoute-visualisation des sons

Nous allons commencer par implémenter une fonction Matlab permettant de visualiser le signal temporel, l'analyse spectrale et le spectrogramme d'un son donné.

On obtient par exemple pour le deuxième signal les tracés suivants :

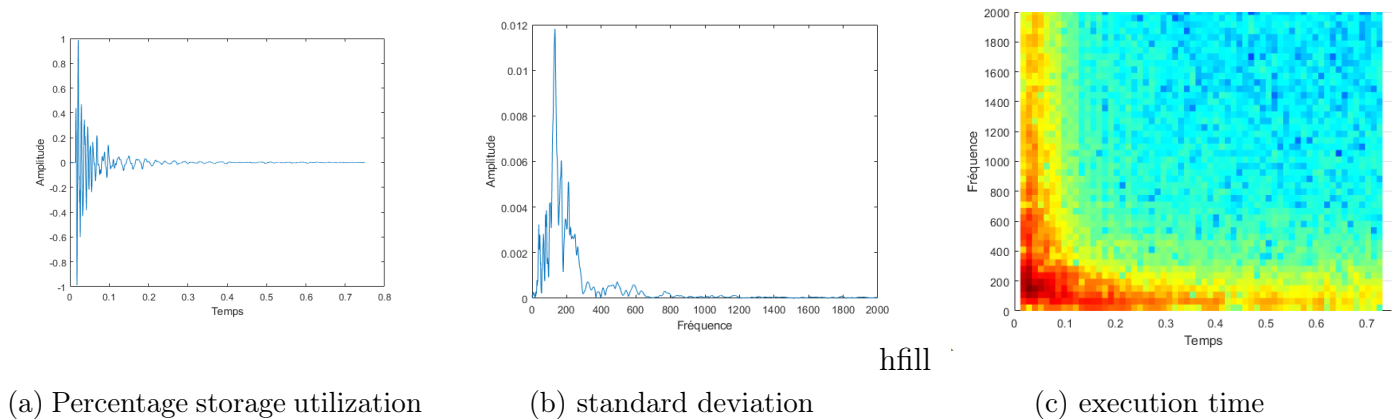
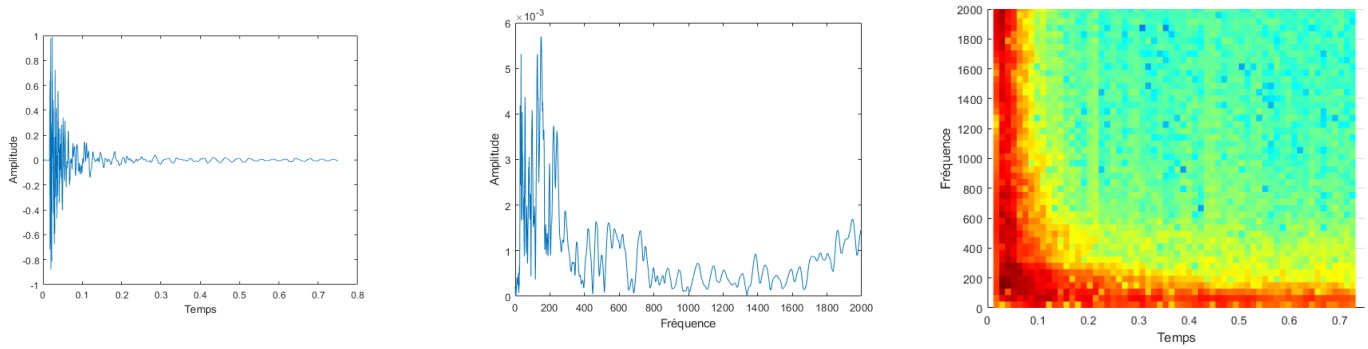


FIGURE 1

On observe alors à partir de ces tracés que :

- L'amplitude du bruit décroît rapidement
- Le signal contient majoritairement des basses fréquences
- le contenu fréquentiel devient moins riche au fil du temps

Pour le 18ème son, on obtient :



(a) Percentage storage utilization

(b) standard deviation

(c) execution time

FIGURE 2

Cette fois, on observe :

- On ne voit pas de différences marquantes entre les allures temporelles
- Ce son contient bien plus de hautes fréquences que le son 2

En effet, en écoutant les deux signaux, le son 2 est bien plus agréable que le son 18.

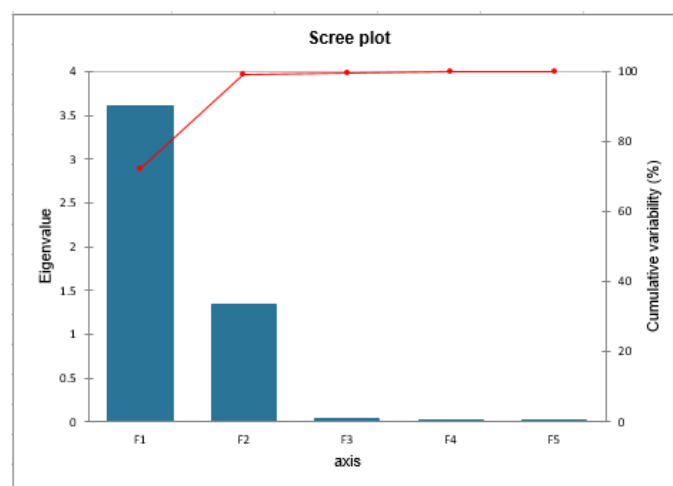
Par ailleurs, on note l'influence de la taille de la fenêtre sur la précision dans les domaines fréquentiel et temporel pour le spectrogramme.

2 Analyse subjective des sons

Pour une analyse subjective des sons, on cherche maintenant à effectuer une ACP afin de réaliser :

- L'analyse des composantes principales
- L'étude des individus

FIGURE 3 – Cumulative variability



On observe ainsi que les deux premiers axes retiennent **99.114 %** d'inertie.

FIGURE 4

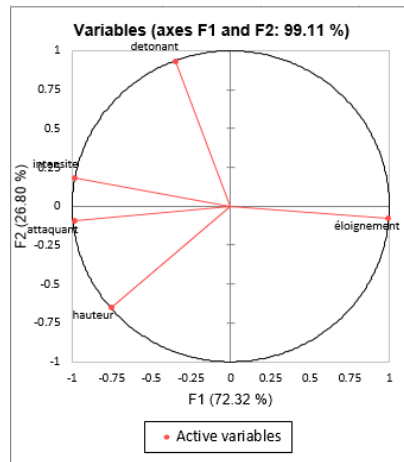


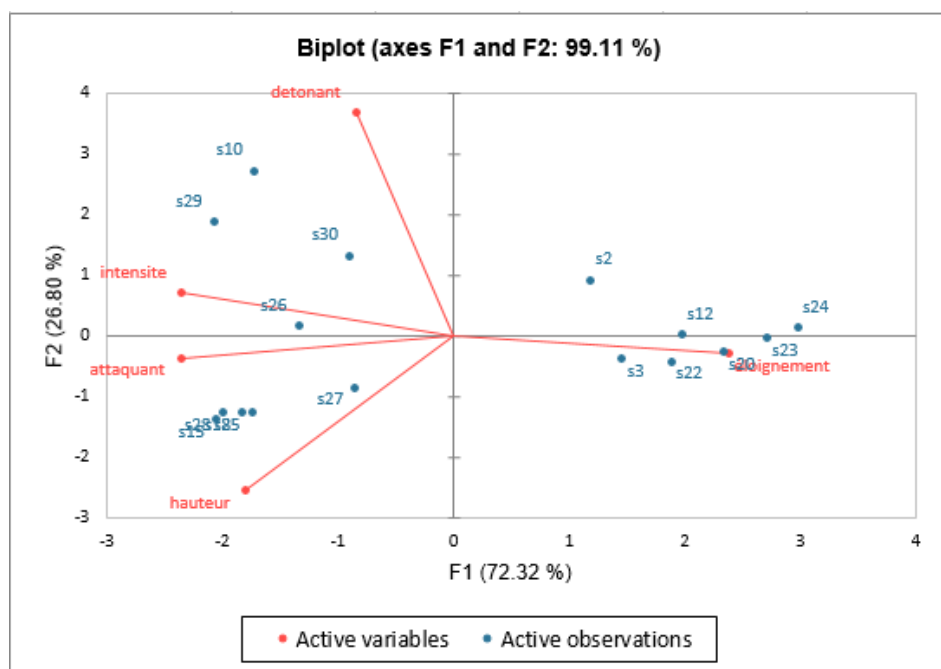
FIGURE 5 – Squared cosines

	F1	F2	F3	F4	F5
hauteur	0.569	0.420	0.009	0.002	0.000
detonnant	0.124	0.873	0.002	0.001	0.000
attaquant	0.972	0.009	0.019	0.001	0.000
éloignement	0.988	0.006	0.000	0.005	0.000
intensite	0.963	0.033	0.001	0.000	0.003

On observe à partir de ces figures que :

- La première composante principale contribue beaucoup à : l'attaquant, l'éloignement et l'intensité
- La deuxième composante principale contribue beaucoup au détonnant
- La hauteur repose sur les deux composantes

FIGURE 6 – biplot



À partir de cette figure, on observe que :

- Un son situé à l'extrémité positive de la CP 1 est caractérisé par un éloignement important
- Un son situé à l'extrémité négative de la CP 1 est caractérisé par une intensité et un attaquant importants et une hauteur assez importante
- Un son situé à l'extrémité négative de la CP 2 est caractérisé par une hauteur assez importante
- Un son situé à l'extrémité positive de la CP 2 est caractérisé par un détonnant important
- Les sons 2, 12, 24, 3, 22, 20 et 23 contribuent le plus à la CP 1
- Pour la CP 2, c'est moins homogène
- Trois clusters seombent se dégager de cette analyse multivariée

Ainsi, on observe que le cluster à droite de la figure est le plus agréable à l'écoute. Les deux clusters à gauche, sont globalement assez désagréables même si on observe que le cluster en haut à gauche reste quand même assez qualitatif.

3 Analyse objective des sons

On cherche à extraire différents descripteurs des sons. On peut ensuite utiliser ces descripteurs pour comparer les sons entre eux.

Par exemple, pour les sons 2 et 18, on obtient les descripteurs suivants :

Son 2 :

```
lrms: 17.6118
lmax: 0.9885
ldb: 57.1380
kurtosis: 58.6986
tmax: 0.0180
tdec: 0.029
pb: [0.0854 0.5467 0.1293 0.0478 0.0347 0.0288 0.0207 0.0157]
dba: [7×1 double]
cdg: 8
tr1: 0.2604
tr2: 0.9695
tr3: 0.0083
```

Son 18 :

```
lrms: 14.6980
lmax: 0.9885
ldb: 55.5671
kurtosis: 39.4787
tmax: 0.0239
tdec: 0.052
pb: [0.1078 0.3457 0.2523 0.2467 0.5667 0.2517 0.2212 0.1761]
dba: [7×1 double]
cdg: 8
tr1: 0.1496
tr2: 0.3389
tr3: 0.6449
```

4 Étude des corrélations

On s'intéresse à l'étude des corrélations (corrélation linéaire de Pearson) entre les descripteurs sensoriel et les descripteurs du signal.

FIGURE 7 – Matrice de corrélation

	hauteur	detonant	attaquant	éloignement	intensité	Kurtosis	tmax	decroissar	pb1	pb2	pb3	pb4	pb5	pb6	pb7	pb8	lrms	tr1
hauteur	1	-0.334	0.792	-0.701	0.625	0.629	-0.253	-0.249	-0.657	-0.593	0.432	0.520	0.770	0.771	0.826	0.864	-0.781	-0.663
detonant	-0.334	1	0.254	-0.418	0.515	-0.214	-0.464	-0.399	0.207	-0.104	-0.327	0.156	-0.218	-0.132	-0.244	-0.341	0.008	0.391
attaquant	0.792	0.254	1	-0.968	0.946	0.488	-0.566	-0.520	-0.478	-0.717	0.225	0.625	0.661	0.706	0.695	0.666	-0.754	-0.406
éloignement	-0.701	-0.418	-0.968	1	-0.988	-0.462	0.634	0.583	0.464	0.672	-0.137	-0.569	-0.596	-0.637	-0.619	-0.568	0.748	0.334
intensité	0.625	0.515	0.946	-0.988	1	0.385	-0.634	-0.569	-0.384	-0.637	0.115	0.568	0.523	0.597	0.568	0.510	-0.680	-0.252
Kurtosis	0.629	-0.214	0.488	-0.462	0.385	1	-0.315	-0.356	-0.670	-0.502	0.477	0.078	0.146	0.177	0.212	0.359	-0.706	-0.490
tmax	-0.253	-0.464	-0.566	0.634	-0.634	-0.315	1	0.586	0.345	0.308	-0.065	-0.128	-0.201	-0.250	-0.147	-0.148	0.603	0.201
tps decrois	-0.249	-0.399	-0.520	0.583	-0.569	-0.356	0.586	1	0.277	0.574	0.354	0.077	-0.051	-0.174	-0.130	-0.279	0.609	-0.039
pb1	-0.657	0.207	-0.478	0.464	-0.384	-0.670	0.345	0.277	1	0.113	-0.218	-0.263	-0.366	-0.312	-0.348	-0.396	0.855	0.859
pb2	-0.593	-0.104	-0.717	0.672	-0.637	-0.502	0.308	0.574	0.113	1	-0.176	-0.280	-0.356	-0.359	-0.370	-0.465	0.538	-0.088
pb3	0.432	-0.327	0.225	-0.137	0.115	0.477	-0.065	0.354	-0.218	-0.176	1	0.394	0.167	0.216	0.200	0.184	-0.188	-0.383
pb4	0.520	0.156	0.625	-0.569	0.568	0.078	-0.128	0.077	-0.263	-0.280	0.394	1	0.589	0.609	0.508	0.466	-0.281	-0.437
pb5	0.770	-0.218	0.661	-0.596	0.523	0.146	-0.201	-0.051	-0.366	-0.356	0.167	0.589	1	0.820	0.878	0.770	-0.464	-0.540
pb6	0.771	-0.132	0.706	-0.637	0.597	0.177	-0.250	-0.174	-0.312	-0.359	0.216	0.609	0.820	1	0.952	0.888	-0.445	-0.525
pb7	0.826	-0.244	0.695	-0.619	0.568	0.212	-0.147	-0.130	-0.348	-0.370	0.200	0.508	0.878	0.952	1	0.894	-0.464	-0.533
pb8	0.864	-0.341	0.666	-0.568	0.510	0.359	-0.148	-0.279	-0.396	-0.465	0.184	0.466	0.770	0.888	0.894	1	-0.534	-0.528
lrms	-0.781	0.008	-0.754	0.748	-0.680	-0.706	0.603	0.609	0.855	0.538	-0.188	-0.281	-0.464	-0.445	-0.464	-0.534	1	0.616
tr1	-0.663	0.391	-0.406	0.334	-0.252	-0.490	0.201	-0.039	0.859	-0.088	-0.383	-0.437	-0.540	-0.525	-0.533	-0.528	0.616	1

On associe à chaque descripteur sensoriel un descripteur du signal qui l'explique :

- Hauteur : pb8
- Détonnant : t_{max} coefficient le plus élevé mais quand même assez faible
- Attaquant : L_{RMS}
- Éloignement : L_{RMS}
- Intensité : L_{RMS}

On note en particulier la forte corrélation entre L_{RMS} et les descripteurs sensoriels différents autre que le détonnant.

5 Étude de la qualité des sons

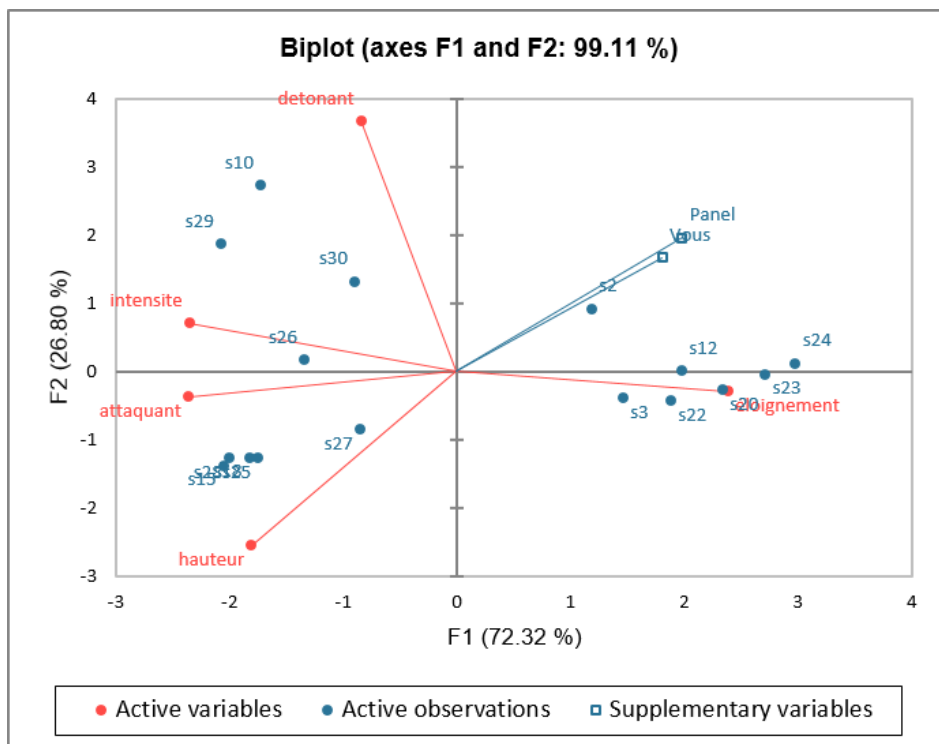
1. On classe les signaux de la façon suivante :

Panel Vous

5.5 6
4.7 4
4.5 4
5.3 7
2.9 2
3.4 3
4.8 6
5.3 6
5.7 7
6.2 7
3.2 2
4.3 4
3.8 5
3.1 3
4.6 5

2. On projette les cotations de qualité en variables supplémentaires sur le plan de l'ACP des descripteurs sensoriels. On obtient la figure suivante :

FIGURE 8 – Projection



On obtient alors des résultats cohérents avec notre préférence pour le cluster de droite. On a les mêmes préférences que le panel.

3. On représente la cotation de qualité en fonction de chaque descripteur. On obtient les figures suivantes :

FIGURE 9 – Intensité

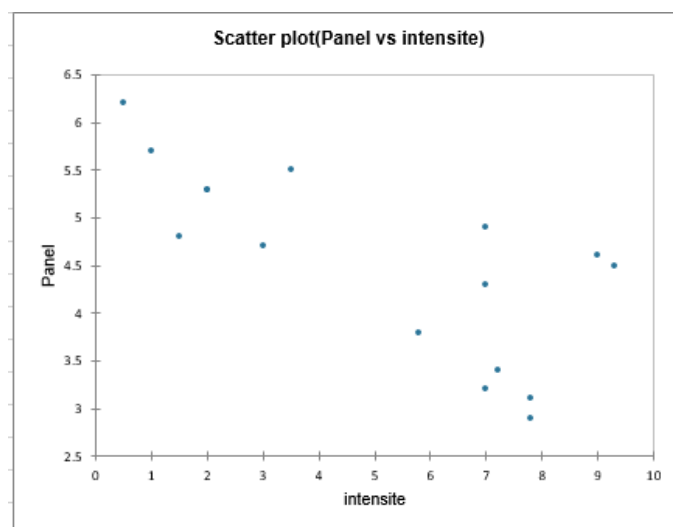


FIGURE 10 – Attaquant

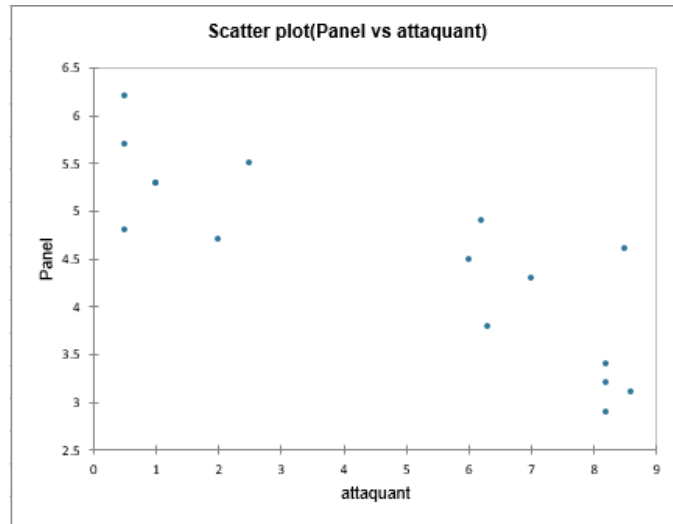


FIGURE 11 – Détonnant

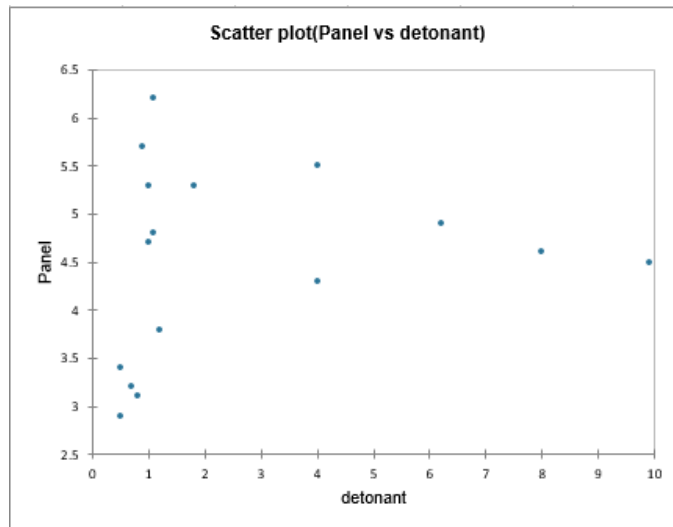


FIGURE 12 – Éloignement

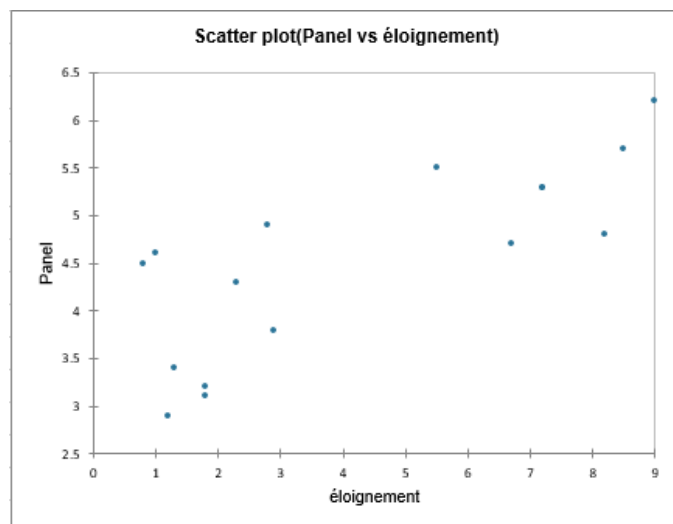
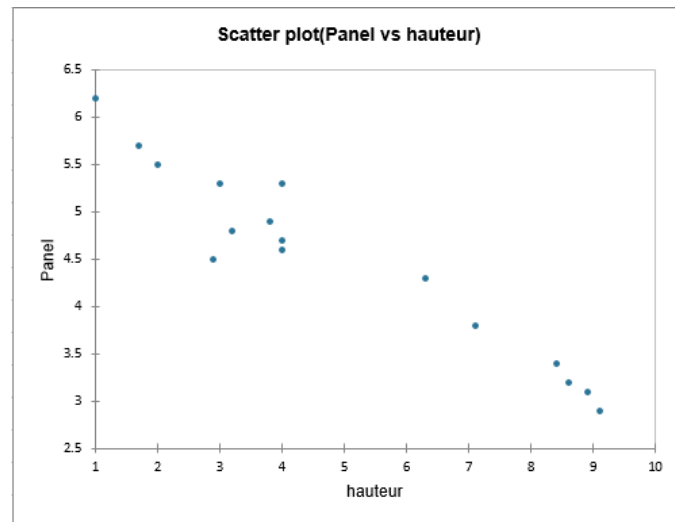


FIGURE 13 – Hauteur



On obtient la matrice de corrélation suivante :

FIGURE 14 – Matrice de corrélation

	hauteur	detonant	attaquant	éloignemen	intensite	Panel	Vous
hauteur	1	-0.334	0.792	-0.701	0.625	-0.961	-0.852
detonant	-0.334	1	0.254	-0.418	0.515	0.169	0.131
attaquant	0.792	0.254	1	-0.968	0.946	-0.837	-0.753
éloigneme	-0.701	-0.418	-0.968	1	-0.988	0.785	0.723
intensite	0.625	0.515	0.946	-0.988	1	-0.732	-0.678
Panel	-0.961	0.169	-0.837	0.785	-0.732	1	0.902
Vous	-0.852	0.131	-0.753	0.723	-0.678	0.902	1

On observe alors que les corrélations entre les cotations de qualité et les descripteurs sont très importantes sauf pour le détonnant.

Enfin, on propose les modèle de régression linéaire suivants :

FIGURE 15 – Panel

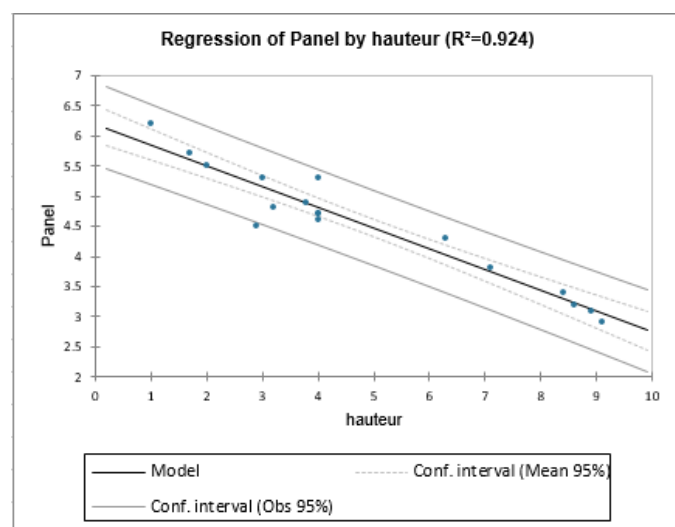
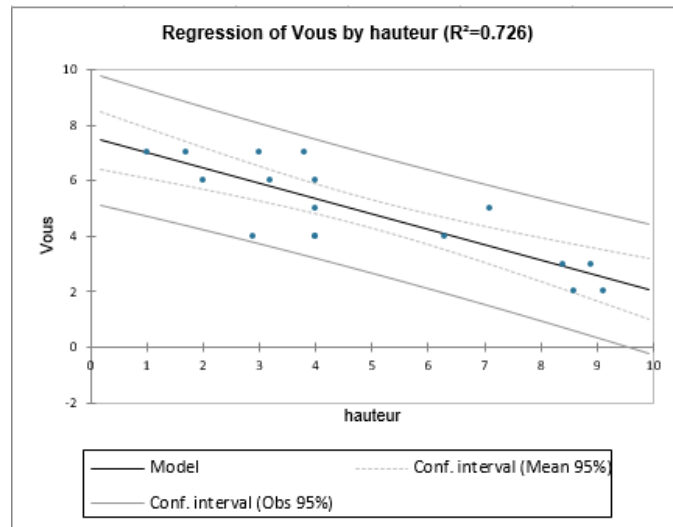


FIGURE 16 – Vous



On obtient $R^2 = 0.924$ pour le cas du panel, une valeur élevée. L'ajustement du modèle est donc correct. La qualité du son décroît donc avec la hauteur. La valeur de R^2 est moins élevée dans le second cas.

4.