

Rapport d'analyse des présences des joueurs aux entraînements

Plan d'expérience

Emmanuel GOHORY
IMT17

26 février 2026

Table des matières

1	Introduction	2
1.1	Contexte de l'expérience	2
1.2	Portée de l'expérience	3
2	Méthode	4
2.1	Collecte des données	4
2.1.1	Nombre de joueurs	4
2.1.2	Météo	4
2.1.3	Match	4
2.2	Démarche	5
3	Résultats	6
3.1	Plan Factoriel	6
3.2	Analyse des variances à un facteur et hypothèses	7
3.2.1	Jour d'entraînement	8
3.2.2	Climat/Temps observé	9
3.2.3	Température	9
3.2.4	Match	10
3.3	Analyse à deux facteurs	10
3.3.1	Météo complète	10
3.3.2	Intempéries et jour d'entraînement	11
3.3.3	Jour d'entraînement et Match	12
3.3.4	Intempéries et préparation de match	13
4	Conclusion	15
5	Sources	16
6	Annexes	17
6.1	Annexe 2 : Code Python	17

Chapitre 1

Introduction

1.1 Contexte de l'expérience

Cette étude s'inscrit dans le cours de *Plan d'expérience* de ma deuxième année d'ingénieur en formation continue à Phelma. Il nous a été demandé de prévoir et mener un plan d'expérience sur le sujet de notre choix.

Aussi, en tant que joueur de rugby à XV pour le club de rugby de Pont de Claix, l'US 2 Ponts, j'ai décidé de mener une étude sur le taux de présence des joueurs aux entraînements en fonction de critères *évidents*. Les joueurs sont tenus d'assister, autant que faire ce peut, à deux entraînements par semaine, le mardi et le jeudi soir (de 19h30 à 21h). Le dimanche un match du championnat dans lequel évolue l'équipe peut être disputé pour classer l'équipe dans sa poule et décider de sa participation, ou non, aux phases finales de la compétition Régionale 1. L'enjeu pour le club, est de se classer le plus haut de participer à ces phases finales, de décrocher un titre et/ou de monter de division pour la saison prochaine afin d'attirer joueurs et sponsors. Il y a donc, un fort besoin de présence aux entraînements afin de structurer une équipe capable de relever ces défis.

Ces faits définis, il apparaît clair que comprendre les facteurs qui peuvent influer sur la quantité de joueurs aux entraînements est, donc, primordial. C'est pourquoi, je décrirais tout au long de ce document, la méthode utilisée, les résultats mesurés et les analyses et/ou conclusions que l'on peut en tirer.

1.2 Portée de l'expérience

Comme implicité dans la partie précédente, cette étude se concentre sur l'effectif senior de l'US 2 Ponts sur la saison actuelle, 2025-2026.

Ainsi, l'étude couvre la période de la pré-saison à aujourd'hui et concerne un effectif total de joueurs avoisinant 70.

En détail :

- Période de mi-aout 2025 à mi-février 2026 à raison de deux entraînements par semaine (sauf période de repos et vacances) soit 47 entraînements
- Effectif total de 74 joueurs

Les critères/facteurs choisient pour ces expériences sont :

- la météo
 - le temps observé
 - la température mesurée
- le jour de l'entraînement
- si un match est prévu en fin de semaine ou si c'est une semaine de repos

Chapitre 2

Méthode

2.1 Collecte des données

2.1.1 Nombre de joueurs

Connaitre le nombre de personne s'étant rendu à chaque entraînement fut simple. En effet, un pointage, tenu par les bénévoles du club, est effectué scrupuleusement pour chaque entraînement. Il permet aux encadrants de l'équipe de connaître le taux de présence de chacun et la situation de l'effectif. Ce document, anonymisé, est disponible à la consultation en *Annexe 1 : Présence aux entraînements 2025-2026*

2.1.2 Météo

Pour la météo, il a suffit de récupérer les données météorologiques enregistrées. Celles-ci sont facilement trouvable sur internet. Pour les jours de temps mitigé, j'ai choisi de considéré qu'il faisait beau si plus de la moitié de la journée (sur sa période d'ensolleillement en fonction de la saison) était dégagée ou nuageuse sans pluie. En effet, bien que l'entraînement ait lieu le soir j'ai considéré que la prise de décision sur une participation à l'entraînement se faisait pendant la journée et non à 19h. Pour la température, j'ai choisi de sélectionner l'extremum dans la journée, pour la même raison. Voir en source (chapitre 5) pour le lien vers le site utilisé.

2.1.3 Match

Le calendrier des matchs est facilement accessiblement via les sites de la fédération française de rugby (lien en chapitre 5). Un match le 1er février fait

considérer les entraînements du jeudi 29 et mardi 27 janvier comme étant des préparations de match ($G = \text{oui}$ dans le plan factoriel, figure 3.1).

2.2 Démarche

Afin d'analyser les résultats obtenus, les expériences seront mises en forme dans un plan factiorel. Dans un second temps, les données de ce plan seront codées.

Les variances à un facteur seront analysés par Excel pour chaque variable. Ainsi nous pourrons rechercher d'éventuels liens direct entre un facteur et le résultat ou exclure les variables qui ne semblent pas pertinentes. A cette occasion nous ferons des tests d'hypothèses. Celles-ci seront définis au début de la partie 3.2

Puis, nous utiliserons un script python afin de généré des graphiques de variances à deux facteurs pour permettre une analyse visuelle des influences des facteurs deux à deux. Ce script python est accessible en *Annexe 2 : 6.1*

Chapitre 3

Résultats

3.1 Plan Factoriel

Les facteurs sont nommés de la façon suivante :

- Jour de la semaine : D
- Climat : W
- Température : T
- Match le week-end : G

On note alors, le tableau du plan factoriel tel que :

trial number	D	W	T	G	Presence moyenne
1 mardi		beau temps	>10°C	oui	32,00
2 mardi		beau temps	>10°C	non	30,20
3 mardi		beau temps	<=10°C	oui	32,00
4 mardi		beau temps	<=10°C	non	29,00
5 mardi		pluie	>10°C	oui	34,00
6 mardi		pluie	>10°C	non	19,00
7 mardi		pluie	<=10°C	oui	27,00
8 mardi		pluie	<=10°C	non	17,50
9 jeudi		beau temps	>10°C	oui	33,71
10 jeudi		beau temps	>10°C	non	29,25
11 jeudi		beau temps	<=10°C	oui	32,00
12 jeudi		beau temps	<=10°C	non	28,00
13 jeudi		pluie	>10°C	oui	34,50
14 jeudi		pluie	>10°C	non	31,00
15 jeudi		pluie	<=10°C	oui	30,75
16 jeudi		pluie	<=10°C	non	29,00

FIGURE 3.1 – Plan factoriel des expériences menées

Ensuite, on peut coder les entrées pour les rendre exploitable pour les calculs et analyses.

Paramètre	valeur	codage
T	>10°C	1
	<=10°C	-1
D	mardi	1
	jeudi	-1
W	beau temps	1
	pluie	-1
G	oui	1
	non	-1

FIGURE 3.2 – Codage des paramètres

On obtient alors le plan factoriel codé suivant :

trial number	D	W	T	G	Présence moyenne
1	1	1	1	1	32,00
2	1	1	1	-1	30,20
3	1	1	-1	1	32,00
4	1	1	-1	-1	29,00
5	1	-1	1	1	34,00
6	1	-1	1	-1	19,00
7	1	-1	-1	1	27,00
8	1	-1	-1	-1	17,50
9	-1	1	1	1	33,71
10	-1	1	1	-1	29,25
11	-1	1	-1	1	32,00
12	-1	1	-1	-1	28,00
13	-1	-1	1	1	34,50
14	-1	-1	1	-1	31,00
15	-1	-1	-1	1	30,75
16	-1	-1	-1	-1	29,00

FIGURE 3.3 – Plan factoriel codé

3.2 Analyse des variances à un facteur et hypothèses

Pour cette partie, nous lancerons des analyses de variances à un facteur sur toutes les variables testées par ces expériences pour essayer de trouver les facteurs les plus impactant sur le taux de présence à l'entraînement.

On vérifiera aussi la valeur P de chaque paramètres pour tester l'hypothèse que le paramètre étudié à un impact direct sur le taux de présence. Il faudra alors, s'assurer que la valeur P (aussi noté *Probabilité* ou *P value*) soit inférieur à 5% pour pouvoir rejeter l'Hypothèse Nulle qui stipule que lien entre paramètre et résultat n'est pas évident. Voici, les quatres hypothèses qui seront testées :

1. H1 : Il existe un lien direct entre la présence à l'entraînement et le jour d'entraînement

2. H2 : Il existe un lien direct entre le temps au jour de l'entraînement et la présence à l'entraînement
3. H3 : Il existe un lien direct entre la température et la présence à l'entraînement
4. H4 : Il existe un lien direct entre la présence ou l'absence d'un match le week-end sur la présence à l'entraînement de cette même semaine.

3.2.1 Jour d'entraînement

Commençons par l'analyse de l'impact du jour d'entraînement sur la présence aux entraînements.

RAPPORT DÉTAILLÉ JOUR DE LA SEMAINE					
Groupes	Nombre d'échantillons	Somme	Moyenne	Variance	
mardi	8	220,7	27,5875	37,8183929	
jeudi	8	248,2142857	31,02678571	5,24808673	
ANALYSE DE VARIANCE					
Source des variations	Somme des carrés	Degré de liberté	Moyenne des carrés	F	Probabilité Valeur critique pour F
Entre Groupes	47,3147449	1	47,3147449	2,19728872	0,16041106 4,600109937
A l'intérieur des groupes	301,4653571	14	21,5332398		
Total	348,780102	15			

FIGURE 3.4 – Analyse des résultats pour le jour d'entraînement

On remarque directement que la présence à l'entraînement est supérieure les jeudis au mardis. De plus, la variance de présence des jeudis est bien plus petite que celle des mardis. Ce qui signifie que, non seulement il y a plus de monde aux entraînements les jeudis mais aussi que la présence est plus stable ces jours en comparaison aux mardis.

En revanche, la valeur P étant supérieur à 5% ($Probabilité = 16,04\%$) on ne peut donc pas rejeter l'Hypothèse nulle ce qui signifie qu'il n'y a pas suffisamment de preuve pour pouvoir affirmer qu'il y ait une différence significative entre le mardi et le jeudi sur le taux de présence aux entraînements. On peut alors dire, qu'il semble avoir un impact du jour de la semaine sur le nombre de joueur aux entraînements mais que cet impact, à lui seul, n'est pas suffisamment significatif.

3.2.2 Climat/Temps observé

RAPPORT DÉTAILLÉ CLIMAT				
Groupes	Nombre d'échantillons	Somme	Moyenne	Variance
beau temps	8	246,1642857	30,77053571	3,806945153
pluie	8	222,75	27,84375	41,12388393
ANALYSE DE VARIANCE				
Source des variations	Somme des carrés	Degré de liberté	Moyenne des carrés	F
Entre Groupes	34,26429847	1	34,26429847	1,525202146
A l'intérieur des groupes	314,5158036	14	22,46541454	
Total	348,780102	15		
Probabilité	Valeur critique pour F			
0,237159782	4,600109937			

FIGURE 3.5 – Analyse des résultats pour le climat/temps observé

De la même manière que pour l'analyse précédente, on observe que la valeur P est supérieure à 5%, ainsi on ne pourra pas conclure d'un lien direct, clair et évident entre le temps du jour de l'entrainement et la présence à l'entrainement. On remarque, tout de même, que la variance part temps pluvieux est très élevée en comparaison des jours de beau temps, ce qui signifie que d'autres facteurs doivent influé sur la présence les jours de pluies tandis que les jours de beaux temps, la présence semble plus stable.

3.2.3 Température

RAPPORT DÉTAILLÉ TEMPERATURE				
Groupes	Nombre d'échantillons	Somme	Moyenne	Variance
>10°C	8	243,6642857	30,45803571	24,99572066
≤10°C	8	225,25	28,15625	21,80245536
ANALYSE DE VARIANCE				
Source des variations	Somme des carrés	Degré de liberté	Moyenne des carrés	F
Entre Groupes	21,1928699	1	21,1928699	0,9057135
A l'intérieur des groupes	327,5872321	14	23,39908801	
Total	348,780102	15		
Probabilité	Valeur critique pour F			
0,357396717	4,600109937			

FIGURE 3.6 – Analyse des résultats pour la température

L'analyse de la température seule ne montre rien de concluant. La valeur P est haute, aucune différence notable entre les moyennes et les variances des deux états de la température. De ce fait, il est impossible de conclure a un impact significatif de la température sur la présence aux entraînements.

3.2.4 Match

RAPPORT DÉTAILLÉ MATCH					
Groupes	Nombre d'échantillons	Somme	Moyenne	Variance	
oui	8	255,9642857	31,99553571	5,678730867	
non	8	212,95	26,61875	27,62709821	
<hr/>					
ANALYSE DE VARIANCE					
Source des variations	Somme des carrés	Degré de liberté	Moyenne des carrés	F	Probabilité Valeur critique pour F
Entre Groupes	115,6392985	1	115,6392985	6,944087666	0,01958817 4,600109937
A l'intérieur des groupes	233,1408036	14	16,65291454		
Total	348,780102	15			

FIGURE 3.7 – Analyse des résultats pour les semaines de match ou non

On voit que **la valeur P est inférieur à 5%** ce qui prouve que l'hypothèse H4 est vérifiée. Il y a donc un lien direct entre la présence aux entraînements et les matchs. Au vu des moyennes, il semblerait que l'on puisse admettre que la présence aux entraînements est plus grande et est plus stable les semaines de match.

3.3 Analyse à deux facteurs

3.3.1 Météo complète

Dans cette partie nous observerons la présence à l'entraînement en fonction des conditions météo complète (température et temps).

Nous avons déterminé, dans la partie précédente (voir partie et) que l'on ne pouvons pas établir de lien entre l'une de ces variables et la présence aux entraînements. Cependant, qu'en est-il d'un lien combiné entre ces deux variables et la présence ?

Cette question nous permet de vérifier si on observer des différences de présence en été ou en hiver. Ainsi, grâce aux données récoltés et au script python on peut générer le graphique suivant :

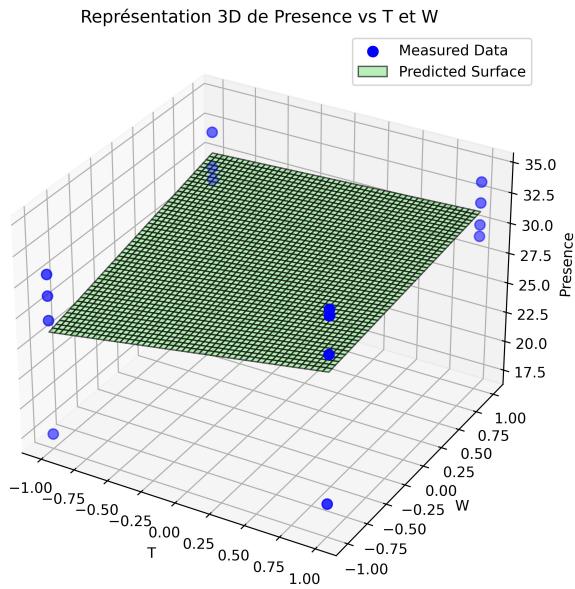


FIGURE 3.8 – Influence des conditions météo T et W sur la présence à l’entraînement

Graphiquement, on voit que, effectivement il semble y avoir plus de monde aux entraînements lorsqu’il fait beau et chaud ($T = 1$, $W = 1$) en comparaison des jours froids et pluvieux ou même simplement pluvieux. Cependant, on note que les points des résultats d’expériences des jours pluvieux ($W = -1$) sont très éparpillés. Ce qui signifie qu’un ou plusieurs autre(s) paramètre(s) doit(vent) influer sur les résultats par temps pluvieux.

3.3.2 Intempéries et jour d’entraînement

On vient de voir que les jours pluvieux donnent des résultats qui semblent influencés par un autre paramètre. C’est pourquoi, nous allons regarder les moyennes de présence en fonction des intempéries (W) et du jour de la semaine (D).

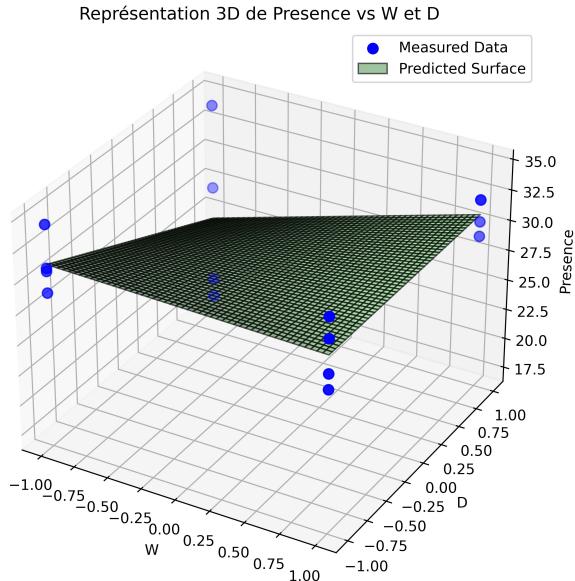


FIGURE 3.9 – Influence des intempéries W et du Jour d'entraînement D sur la présence à l'entraînement

On voit que les résultats de moyenne de présence sont relativement stables les jours de beaux temps ($W = 1$) mais les mardis pluvieux présentent des résultats très variés avec notamment 2 expériences très faibles, et une correcte (aux alentours de 32 personnes)

3.3.3 Jour d'entraînement et Match

Nous allons maintenant nous intéresser à la présence aux entraînements en fonction des semaines de match (G) et du jour de la semaine (D). L'intérêt est d'observer si les joueurs sont plus ou moins présents pour préparer les matchs et si il y a une différence entre un jeudi et un mardi.

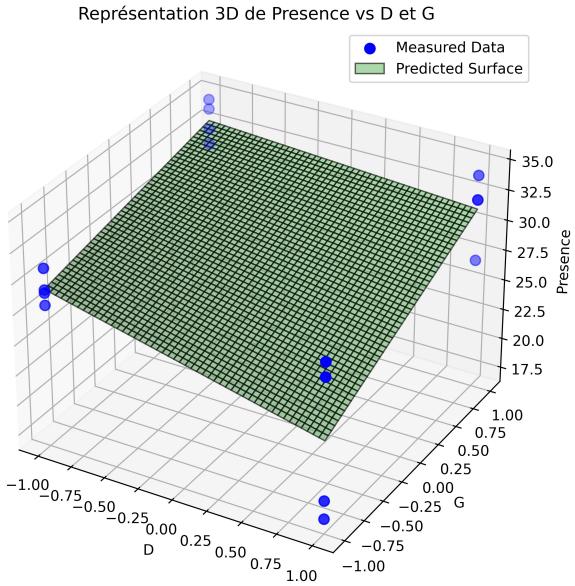


FIGURE 3.10 – Influence du Jour d’entraînement D et du Match le week-end G sur la présence à l’entraînement

On pouvait s’y attendre mais on remarque que, en moyenne, les mardis sans matchs sont délaissés par les joueurs tandis que les jeudis d’avant matchs sont plébiscités. Observation plus surprenante, il apparaît que les jeudis sans matchs comptent pratiquement autant de joueurs que les mardis de préparation de match et ce, avec la même stabilité (dispersions des points similaires).

3.3.4 Intempéries et préparation de match

Nous venons de voir que les entraînements du mardi hors préparation de match semblaient délaissés (en moyenne). On a pu aussi constater que les intempéries semblaient impactés la présence sans pouvoir distinguer de lien évident mais que les entraînements des mardis pluvieux semblaient moins intéresser les joueurs que les autres jours sous d’autres conditions. Nous allons alors, désormais voir l’influence d’une préparation de match (G) et des intempéries (W) sur la présence aux entraînements.

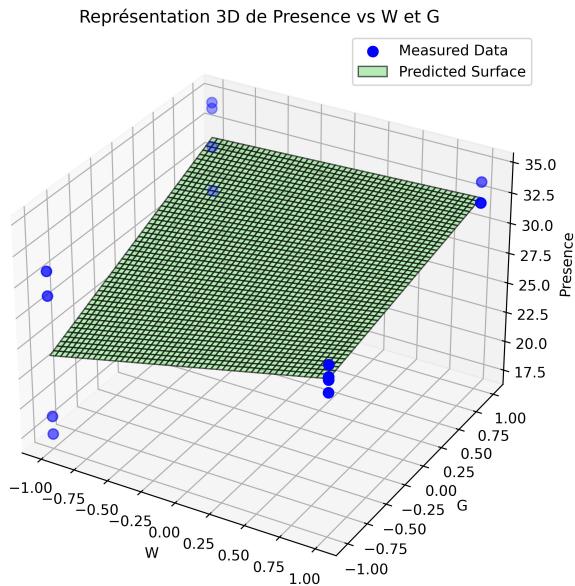


FIGURE 3.11 – Influence du Climat/Intempéries W et du Match le week-end G sur la présence à l'entraînement

On voit que la combinaison du mauvais temps ($W = -1$) et de l'absence de jour de match ($G = -1$) présente les pires résultats moyens avec une variance élevée. Les trois autres combinaisons donnent des résultats similaires avec une variance un peu plus élevée pour le mauvais temps.

Il semble donc correct de dire que les intempéries, cumulées avec d'autres facteurs est un facteur déterminant sur la présence à l'entraînement.

Chapitre 4

Conclusion

Grâce aux résultats de ce plan d'expérience, on a pu démontrer que seul un match le week-end est déterminant sur la présence à l'entraînement en tant que paramètre seul. Cependant, on a pu montrer que les intempéries sont un facteurs non déterminant en soi mais aggravant en combinaison à d'autres facteurs tel que le jour d'entraînement, si ce sont des préparations de match ou non ou s'il fait froid.

On doit néanmoins, prendre du recul sur cette étude car les seuils choisis pour le temps et la météo peuvent faire l'objet de débats. Enfin, et pour clore cette étude, on peut se questionner sur les facteurs hors de portée de cette étude. On peut, notamment, penser à l'influence des vacances scolaires, de l'avancé dans la saison, des présences en fonction des postes de chaque joueur ou bien des résultats de l'équipe. Dans une prochaine étude sur le même sujet, il pourrait aussi être intéressant d'étudier ces phénomènes avec d'autres équipes, de les comparer et d'augmenter les échantillons et expériences ou bien même de faire une comparaison avec l'année précédente, lors de laquelle des primes de matchs étaient versées aux joueurs en fonction de leurs présences aux entraînements mais où les résultats sportifs étaient moins probants.

Chapitre 5

Sources

1. Relevés Météorologiques : fr.weatherspark.com
2. FFR : US 2 Ponts - Calendrier
3. Github : projet python de génération des graphiques ANOVA 2

Chapitre 6

Annexes

6.1 Annexe 2 : Code Python

des trucs

```
1 import numpy as np
2 import pandas as pd
3 import matplotlib.pyplot as plt
4 from mpl_toolkits.mplot3d import Axes3D
5 import statsmodels.api as sm
6 from statsmodels.formula.api import ols
7
8 '''
9 P : Numero d'experience
10 D : Jour d'entraînement, mardi ou jeudi (1 ou -1)
11 W : Meteo, beau temps ou pluie (1 ou -1)
12 T : Temperature pendant l'entraînement, chaud (>10 C) ou
     froid (<=10 C) (1 ou -1)
13 G : Semaine de match ? Match le dimanche ou repos le dimanche
     (1 ou -1)
14 Presence : Moyenne des joueurs presents a l'entraînement sur
     ces jours-la
15 '''
16 # Lecture du fichier Excel contenant les donnees de presence
17 df = pd.read_excel("data_releves_presence.xlsx")
18
19 print(df)
20 # Definition de la formule du modèle linéaire avec les
     interactions
21 formula = 'Presence ~ D + W + T + G + D:W + D:T + D:G + W:T +
     W:G + T:G'
22
23 # Fit model
24 model = ols(formula, data=df).fit()
```

```

25
26
27 # Creation du graph 3D
28 def plot_3d_predictions(model, df, factor_x, factor_y,
29     response_var, num_points=50):
30     fig = plt.figure(figsize=(10, 7))
31     ax = fig.add_subplot(111, projection='3d')
32
33     # Scatter plot of the measured data
34     ax.scatter(df[factor_x], df[factor_y], df[response_var],
35                color='blue', label='Measured Data', s=50)
36
37     # Generate a grid of values for Temperature and
38     # Stirring_rate
39     x_range = np.linspace(df[factor_x].min(), df[factor_x].
40                           max(), num_points)
41     y_range = np.linspace(df[factor_y].min(), df[factor_y].
42                           max(), num_points)
43     X, Y = np.meshgrid(x_range, y_range)
44
45     # Create a DataFrame for the grid and set other factors
46     # to their mean values
47     grid_data = pd.DataFrame({factor_x: X.ravel(), factor_y:
48                               Y.ravel()})
49     for col in df.columns:
50         if col not in [factor_x, factor_y, response_var]:
51             grid_data[col] = df[col].mean()
52
53     # Predict filtration rate on the grid using the model
54     Z = model.predict(grid_data).values.reshape(num_points,
55                                                   num_points)
56
57     # Plot the predicted surface
58     ax.plot_surface(X, Y, Z, color='lightgreen', alpha=0.6,
59                     edgecolor='k', label='Predicted Surface')
60
61     # Labels and legend
62     ax.set_xlabel(factor_x)
63     ax.set_ylabel(factor_y)
64     ax.set_zlabel(response_var)
65     ax.set_title(f"Representation 3D de {response_var} vs {
66     factor_x} et {factor_y}")
67     ax.legend()
68     plt.savefig(f"report/plots/3d_plot_{factor_x}_{factor_y}.
69                 jpeg", dpi=600)
70
71 # Appel de la fonction pour chaque paire de facteurs
72 factor_list = ['T', 'W', 'D', 'G']
73 for i, factor_x in enumerate(factor_list):

```

```
63     for j, factor_y in enumerate(factor_list):
64         if i < j:
65             plot_3d_predictions(model, df, factor_x, factor_y
, 'Presence')
```