

MASTER 2

Le goût et l'innovation en agro-industries

Devoir surveillé de planification expérimentale

1^{ère} partie : Théorie (1h30)

Consignes

- Rédiger le devoir sur feuilles d'examen garantissant l'anonymat
- Il est interdit d'allumer les ordinateurs et calculatrices, ainsi que de consulter le cours.

1- Indiquer la ou les bonnes réponses (temps estimé 20 mn)

Barème : une réponse juste indiquée : + 1pt ; une réponse fautive entourée : - 1 pt
Ce barème vise à décourager de répondre au hasard. Si vous ne connaissez pas la / les bonnes réponses, il vaut mieux ne pas répondre.

1.1 Un plan d'expériences

A – comporte généralement moins d'essais qu'une démarche classique, où ceux-ci sont définis et réalisés séquentiellement en fonction des résultats précédents.

B – est une séquence d'essais pour lesquels les facteurs dont on cherche à quantifier l'influence prennent des valeurs bien définies. ICI

C – permet de minimiser le nombre d'essais pour obtenir un résultat acceptable.

D – permet d'optimiser une séquence d'essais pour quantifier l'influence des facteurs étudiés. ICI

1.2 Le domaine expérimental

A – se définit uniquement par rapport aux valeurs prises par les facteurs dans un plan d'expériences. ICI peut être

B – se définit également par rapport aux valeurs prises par les réponses étudiées dans un plan d'expériences.

C – se définit indépendamment des valeurs prises par les facteurs dans un plan d'expériences.

D – se définit indépendamment des valeurs prises par les réponses étudiées dans un plan d'expériences.

1.3 Un facteur de plan d'expériences

A – est une variable quantitative exclusivement.

B – est une variable qualitative exclusivement.

C – est une variable quantitative ou qualitative. ici

D – est une variable quantitative et qualitative.

1.4 Une réponse de plan d'expériences

A – est une variable quantitative exclusivement. ici

B – est une variable qualitative exclusivement.

C – est une variable quantitative ou qualitative.

D – est une variable quantitative et qualitative.

1.5 Les valeurs prises par des facteurs qualitatifs dans un plan d'expériences s'appellent :

- A – des niveaux. >> quand quantitatif
- B – des paramètres.
- C – **des modalités.>> quand qualitatif** (donc réponse ici)
- D – des variables.

1.6 Il y a interaction entre les facteurs X_1 et X_2 si :

- A – L'effet de X_1 sur la réponse est différent de l'effet de X_2
- B – L'effet de X_1 sur la réponse est inverse de l'effet de X_2
- C – **L'effet de X_1 sur la réponse n'est pas le même suivant la valeur de X_2** ICI maybe
- D – **L'effet de X_2 sur la réponse n'est pas le même suivant la valeur de X_1** ICI maybe

1.7 Dans un plan factoriel complet à deux facteurs et deux niveaux, il y a orthogonalité (Peut être signifie, homogénéité des variance) des facteurs X_1 et X_2 si :

- A – Le coefficient de corrélation entre les valeurs prises par X_1 et les valeurs prises par X_2 est nul pour l'ensemble des essais du plan d'expériences, points au centre inclus.
- B – **Le coefficient de corrélation entre les valeurs prises par X_1 et les valeurs prises par X_2 est nul pour l'ensemble des essais du plan d'expériences en dehors des points au centre.** (MAYBE) (0)
- C – Le coefficient de corrélation entre les valeurs prises par X_1 et les valeurs prises par X_2 est égal à 1 pour l'ensemble des essais du plan d'expériences, points au centre inclus.
- D – Le coefficient de corrélation entre les valeurs prises par X_1 et les valeurs prises par X_2 est égal à 1 pour l'ensemble des essais du plan d'expériences en dehors des points au centre.

1.8 Si une réponse Y présente un optimum dans le domaine expérimental, un plan d'expériences qui pourra le déterminer sera :

- A – **un plan factoriel complet à 2 niveaux.**
- B – **un plan factoriel complet à 3 niveaux.**
- C – **un plan composite centré.**
- D – aucun des trois plans précédents.

1.9 Le traitement statistique permettant d'étudier l'influence conjointe de deux facteurs sur une réponse dans un plan factoriel complet à deux facteurs et trois niveaux peut très bien être :

- A – Une régression linéaire simple
- B – **Une régression linéaire multiple**
- C – Une analyse de variance à un facteur
- D – Une analyse de variance à deux facteurs sans interaction
- E – **Une analyse de variance à deux facteurs et interaction**
- F – Une régression polynomiale

1.10 Quel(s) type(s) de modèle(s) reliant une réponse Y aux facteurs X₁ et X₂ peu(ven)t être établi(s) en traitant un plan composite centré ?

A- $Y = a_1 X_1 + a_2 X_2 + a_0$

B- $Y = a_1 X_1 + a_2 X_2 + a_{12} X_1.X_2 + a_0$

C- $Y = a_1 X_1 + a_2 X_2 + a_{12} X_1.X_2 + b_1 X_1^2 + b_2 X_2^2 + a_0$

D- $Y = a_1 X_1 + a_2 X_2 + a_{12} X_1.X_2 + b_1 X_1^2 + b_2 X_2^2 + b_{12} X_1.X_1^2 + b_{21} X_2.X_2^2 + a_0$

2- Répondre aux questions suivantes (temps estimé : 25 mn)

2.1 Quelle est la matrice expérimentale d'un plan factoriel complet à deux facteurs ayant deux niveaux ?

A	B
0	0
1	1
1	-1
-1	1
-1	-1

- Quel(s) modèle(s) permet de tester ce plan ? Avec quel(s) traitement(s) statistique(s) ? **plan factoriel à deux niveau anova ou regression linéaire**

- Quelle est la bonne manière d'exploiter les points au centre ? (validation de model)

2.2 - Quelle est la matrice expérimentale d'un plan composite centré à deux facteurs ?

A	B
0	0
1	1
1	-1
-1	1
-1	-1
a	A
a	-A
-a	A
-a	-a

- Quel(s) modèle(s) permet de tester ce plan ? **composite centré quadratique**

- Avec quel(s) traitement(s) statistique(s) ? S'il faut deux journées pour réaliser les essais, lesquels doivent être regroupés lors d'une même journée , et pourquoi ? **plan factoriel à deux niveau anova ou regression linéaire**

2.3 X est une variable qualitative. On étudie les variations d'une variable réponse Y en fonction des modalités de X.

- Comment s'appelle la table ci-dessous ? **T'ableau d'analyse des variance**

- Combien de modalités la variable X comporte t-elle ? **p**

- Le nombre d'échantillons analysés étant ici identique pour chaque modalité de la variable X , exprimez ce nombre en fonction de n et p. **n^p**

- Que vaut l'écart type résiduel ? Que représente t-il ? **$\sum_{j=1}^p \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j)^2 = SR$**

somme des carrés résiduelles

- Quelle est la condition pour conclure que la variable X a un effet significatif sur la variable Y avec un risque inférieur à α de se tromper ? pas dans les bornes de fisher

Effet	SC Somme des carrés	ddl Degrés de liberté	CM Carrés moyen	Fischer	Fcrit Confiance $1-\alpha$
Variable X	SX	$p-1$	$SX/p-1$	$\frac{SX/p-1}{SR/n-p}$	Fcrit(p-1, n-p)
Erreur	SR	$n-p$	$SR/n-p$		
Total	ST	$n-1$	$ST/n-1$		

- 2.4 X_1 et X_2 sont deux variables qualitatives, Y une variable réponse. On effectue une analyse de variance de Y à deux facteurs et interaction.
- Quelles sont toutes les conditions qui doivent être en toute rigueur vérifiées pour que ce traitement soit permis ?
 - Que doit-on faire si ces conditions sont loin d'être respectées ?

3- Interprétation d'un plan d'expériences (temps estimé : 45 mn dont 25mn de lecture)

On réalise pour des besoins de recherche un simulacre de yaourt à base de caséines et de protéines sériques en poudre (protéines laitières), gélifié avec de la GDL (glucono delta lactone) à une valeur pH de 4.4 suivant un protocole de fabrication défini.

On veut déterminer, en fonction de la quantité de caséines et protéines sériques du simulacre de yaourt la quantité de GDL à ajouter de manière à obtenir un pH de 4.4.

On construit pour cela un plan d'expériences à 3 facteurs (protéines sériques, caséines, et GDL), avec pour réponse le pH.

Le domaine de variation des facteurs est le suivant :

- Caséines : 2,4 à 5,4%
- Protéines sériques : 0,4 à 2,4%
- GDL : de 1 à 3%

On suppose que l'on peut avoir un modèle quadratique avec le pH, et on décide donc de réaliser un plan composite centré en choisissant α pour l'orthogonalité (1.287), avec 3 points au centre. On obtient la matrice d'essais suivante :

N° Essai	Code	Caséines*	protéines sériques*	GDL*	Caséines	protéines sériques	GDL
1	---	-1	-1	-1	2,73	0,62	1,22
2	"000	0	0	0	3,9	1,4	2
3	a00	-1,287	0	0	2,4	1,4	2
4	+++	1	1	1	5,07	2,18	2,78
5	00A	0	0	1,287	3,9	1,4	3
6	0A0	0	1,287	0	3,9	2,4	2
7	--+	-1	1	1	2,73	2,18	2,78
8	++-	1	1	-1	5,07	2,18	1,22
9	0A0	0	-1,287	0	3,9	0,4	2
10	"000	0	0	0	3,9	1,4	2
11	+--	1	-1	-1	5,07	0,62	1,22
12	00A	0	0	-1,287	3,9	1,4	1
13	--+	-1	1	-1	2,73	2,18	1,22
14	++-	1	-1	1	5,07	0,62	2,78
15	"000	0	0	0	3,9	1,4	2
16	a00	1,287	0	0	5,4	1,4	2
17	--+	-1	-1	1	2,73	0,62	2,78

On a réalisé les essais du plan d'expériences en une seule journée. On a employé deux protocoles de mesure du pH, le 1^{er} consistant à mesurer le pH d'un premier échantillon de yaourt sans le mélanger préalablement (pH set yoghurt), le 2^{ème} consistant à mélanger un deuxième échantillon avant la mesure de pH (pH stirred yoghurt).

On obtient la matrice de résultats suivante :

N° d'essai	pH - set yoghurt	pH - stirred yoghurt
1	4,27	4,31
2	4,09	4,04
3	3,88	3,82
4	4,28	3,8
5	3,81	3,46
6	4,26	4,11
7	3,78	3,67
8	4,59	4,65
9	4	3,85
10	3,94	4
11	4,27	4,44
12	4,91	4,6
13	4,47	4,33
14	3,93	3,73
15	4,19	3,98
16	4,12	4,14
17	3,83	3,45

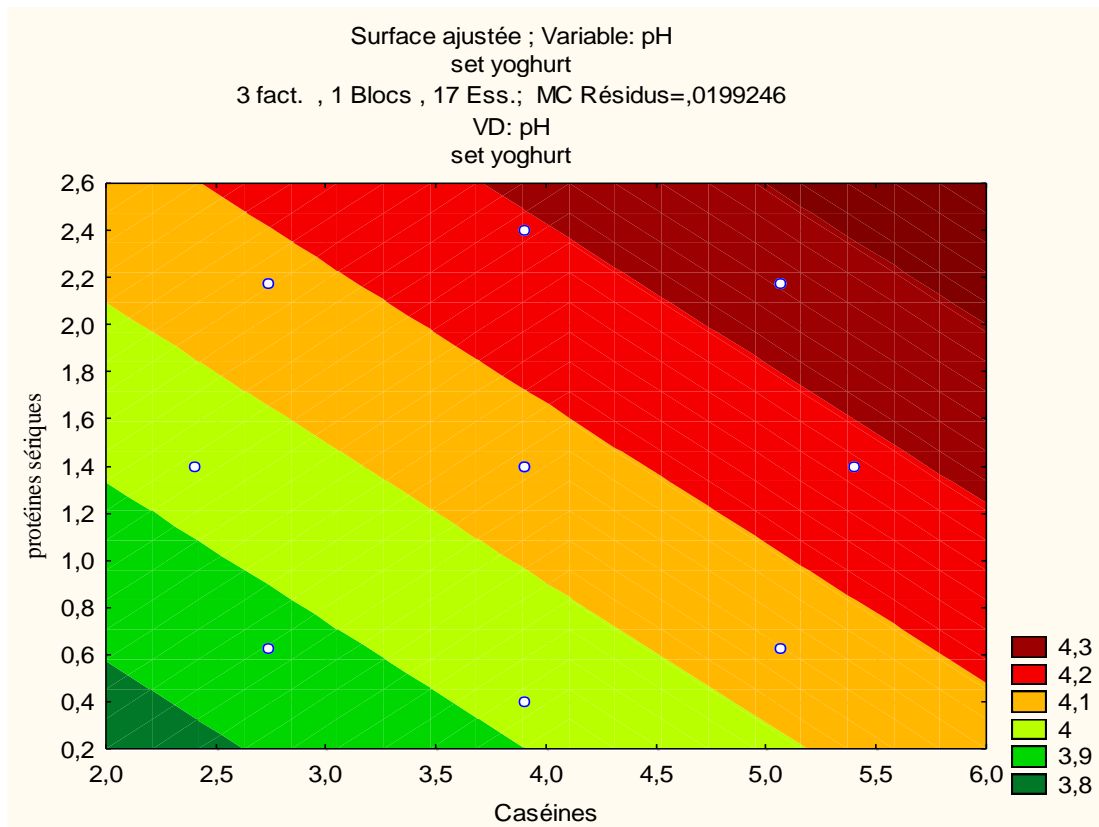
Avec un logiciel de statistique (statistica), on réalise le traitement reliant les deux réponses de pH aux facteurs. On simplifie les modèles obtenus en ne conservant que les coefficients significatifs.

Réponse pH set yoghurt

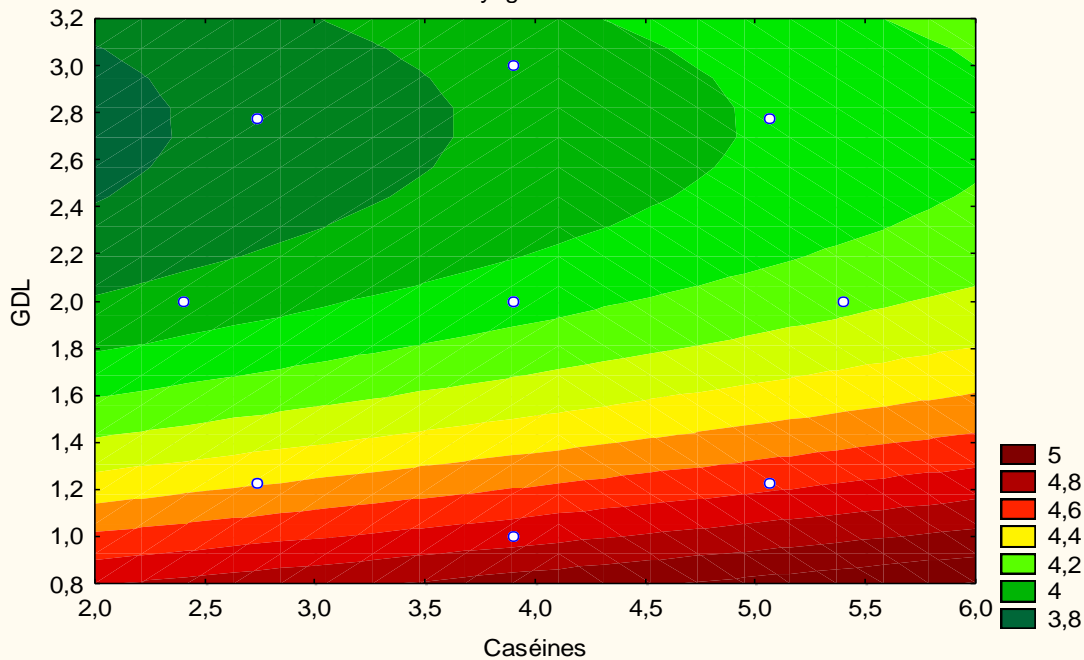
ANOVA; Var.pH set yoghurt R ² =,83845; Aj.: ,7846 (Exercice 2 dans Exercice 3 fact. , 1 Blocs , 17 Ess; MC Résidus=,0199246 VD: pH set yoghurt					
Fact.	SC	dl	MC	F	p
(1)Caséine(L)	0,09357	1	0,09357	4,6964	0,05105
(2)protéines sériques (L)	0,11784	1	0,11784	5,9145	0,03161
(3)GDL(L)	0,90274	1	0,90274	45,3080	0,00002
GDL(Q)	0,12675	1	0,12675	6,3616	0,02679
Erreur	0,23909	12	0,01992		
Total SC	1,48001	16			

Coeffs Régression; R ² =,83845; Aj.:,7846 (Exercice 2 dans Exercic 3 fact. , 1 Blocs , 17 Ess.; MC Résidu=,0199246 VD: pH set yoghurt						
Fact.	Régressn (coeffs)	Err-Type	t(12)	p	-95,% Lim.Conf	+95,% Lim.Conf
Moy/Ord.Orig	5,2622	0,39629	13,2785	0,00000	4,3987	6,12570
(1)Caséines(L)	0,0780	0,03600	2,1671	0,05105	-0,0004	0,15648
(2)protéines sériques (L)	0,1313	0,05401	2,4319	0,03161	0,0136	0,24903
(3)GDL(L)	-1,3298	0,38690	-3,4372	0,00492	-2,1728	-0,48687
GDL(Q)	0,2415	0,09577	2,5222	0,02679	0,0328	0,45025

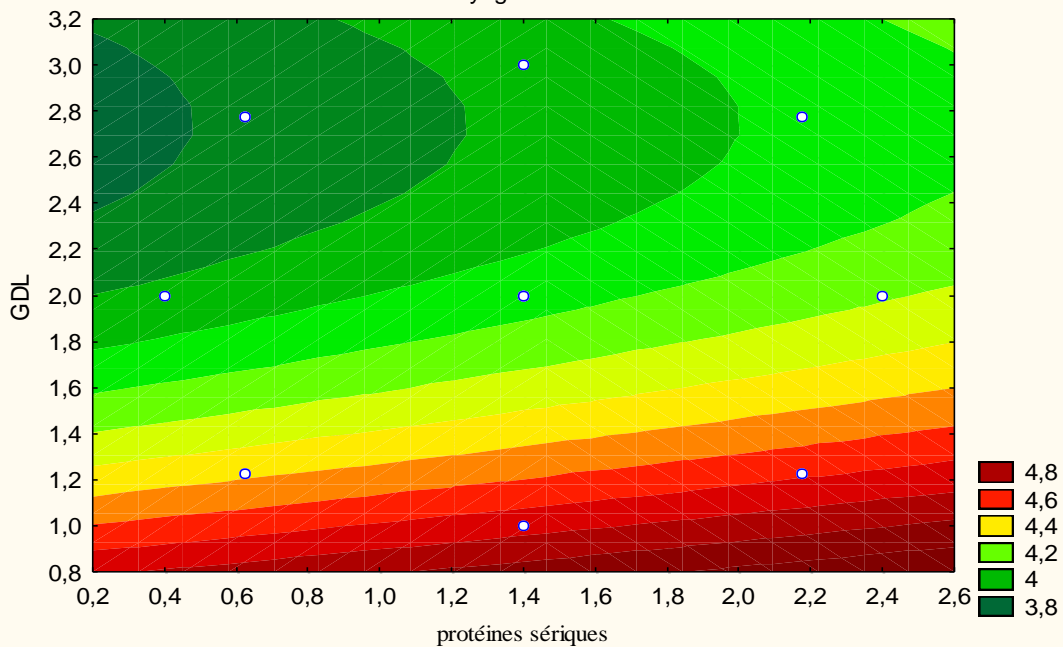
* Les coefficients sont donnés en fonction des facteurs non codés

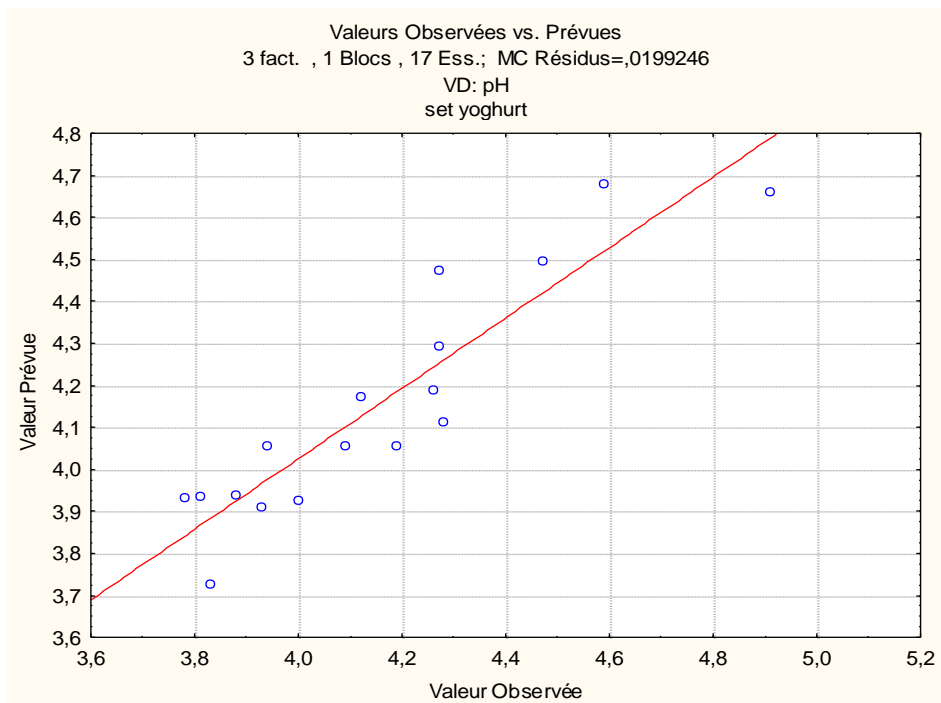


Surface ajustée ; Variable: pH
set yoghurt
3 fact. , 1 Blocs , 17 Ess.; MC Résidus=,0199246
VD: pH
set yoghurt



Surface ajustée ; Variable: pH
set yoghurt
3 fact. , 1 Blocs , 17 Ess.; MC Résidus=,0199246
VD: pH
set yoghurt





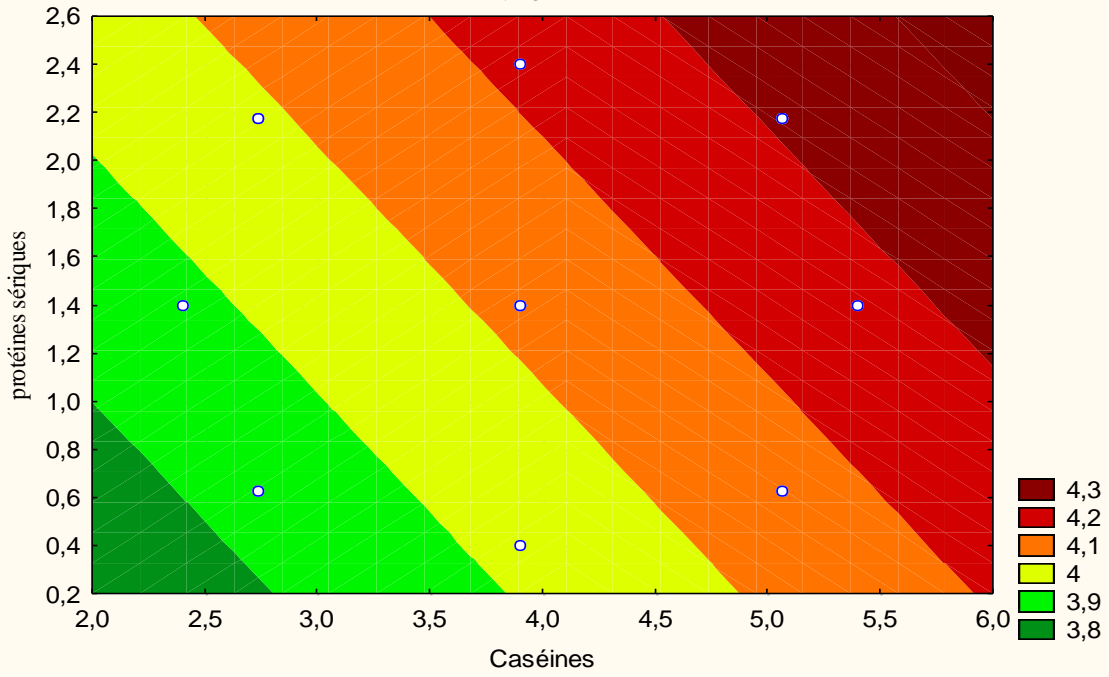
Réponse pH stirred yoghurt

ANOVA; Var.pH stirred yoghurt; R ² =,97987; Aj.:,97523 (Exercice 2 dans Exercice 3 fact. , 1 Blocs , 17 Ess.; MC Résidus=,003216 VD: pH stirred yoghurt					
Fact.	SC	dl	MC	F	p
(1)Caséines(L)	0,14298	1	0,14298	44,461	0,00001
(2)protéines sériques (L)	0,06456	1	0,06456	20,075	0,00061
(3)GDL(L)	1,82774	1	1,82774	568,337	0,00000
Erreur	0,04180	13	0,00321		
Total SC	2,07710	16			

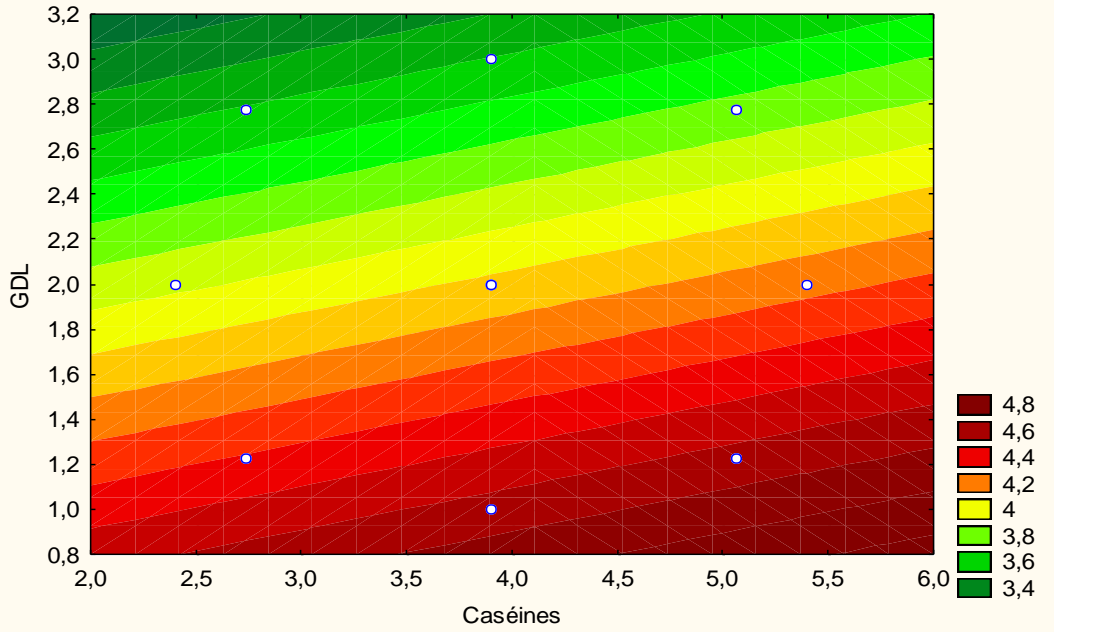
Coeffs Régression; R ² =,97987; Aj.:,97523 (Exercice 2 dans Exercice 3 fact. , 1 Blocs , 17 Ess.; MC Résidus=,003216 VD: pH stirred yoghurt						
Fact.	Régressn (coeffs)	Err-Type	t(13)	p	-95,% Lim.Conf	+95,% Lim.Conf
Moy/Ord.Orig	4,54466	0,07860	57,817	0,00000	4,37484	4,71447
(1)Caséines(L)	0,09646	0,01446	6,668	0,00001	0,06520	0,12771
(2)protéines sériques (L)	0,09722	0,02169	4,480	0,00061	0,05034	0,14410
(3)GDL(L)	-0,51731	0,02169	-23,839	0,00000	-0,56419	-0,47043

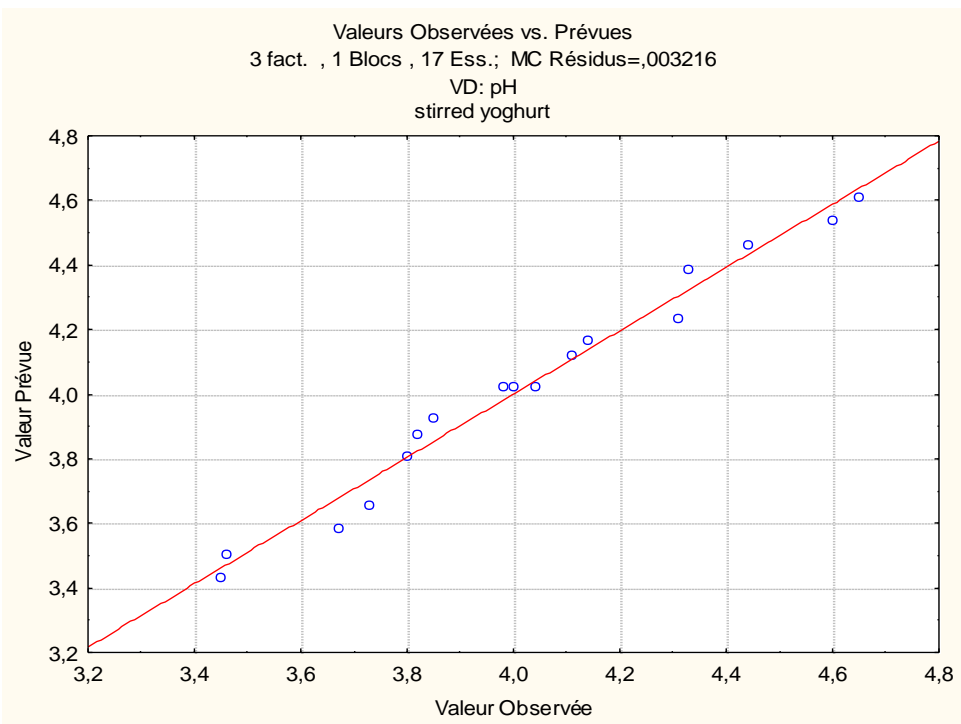
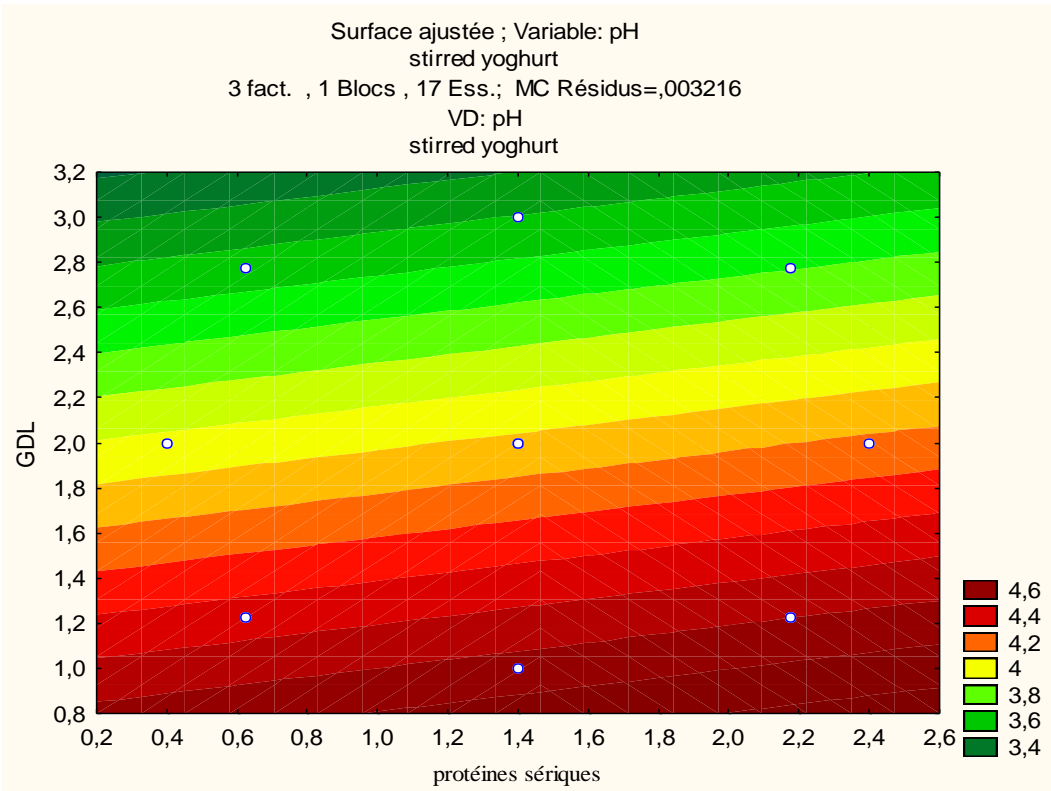
* Les coefficients sont donnés en fonction des facteurs non codés

Surface ajustée ; Variable: pH
 stirred yoghurt
 3 fact. , 1 Blocs , 17 Ess.; MC Résidus=,003216
 VD: pH
 stirred yoghurt



Surface ajustée ; Variable: pH
 stirred yoghurt
 3 fact. , 1 Blocs , 17 Ess.; MC Résidus=,003216
 VD: pH
 stirred yoghurt





Répondez aux questions suivantes :

1- Comment s'appellent le traitement mathématique et les représentations graphiques effectués pour quantifier l'effet des facteurs sur les deux réponses ?

2- Qualifier les modèles retenus pour les deux réponses.

3- Exprimer la valeur de l'écart type résiduel d'estimation du pH avec les données disponibles pour les deux réponses, sans effectuer le calcul.

4- Quel est le meilleur protocole de mesure (stirred ou set yoghurt) ? Justifier la réponse. Le résultat est-il logique ?

On utilise maintenant le modèle du meilleur protocole de mesure pour ajuster la quantité de GDL d'une formule contenant 4% de caséines et 1% de protéines sériques, de manière à obtenir un pH de 4.4.

5- Exprimer la quantité de GDL nécessaire en fonction des trois paramètres teneur en caséine, teneur en protéines sériques, et pH, puis exprimez ce calcul avec les valeurs numériques de ces paramètres, sans l'effectuer (les calculatrices étant interdites).

6- Exprimer mathématiquement une approximation de l'encadrement à 95% de confiance du pH qui sera obtenu en fonction de l'écart type résiduel d'estimation du pH.