Catalogue de modèles non linéaires

référence: Nonlinear Regression with Built-In Models p.101-122

JMP® 12 Specialized Models. Cary, NC: SAS Institute Inc., March 2015

Dans les sciences biologiques et les sciences de la vie, il existe des familles de fonctions de forme connue qui décrivent la relation entre 2 variables. Par exemple, les chercheurs en pharmacologie conduisent des expériences pour comprendre la force de la réponse en fonction de la concentration du médicament. La famille de fonctions à 1 ou 2 compartiments est souvent employé pour l'étude de ce processus. Un autre exemple est l'utilisation des courbes de croissance exponentielle pour prédire le nombre de personnes d'une population en fonction du temps.

Description générale des fonctions

Courbes en forme de S croissantes ou décroissantes (sigmoïde)

- **Fonctions logistiques, fonctions Gompertz ; ces fonctions ont une asymptote supérieure et une asymptote inférieure.**
- **Fonctions logistiques symétriques avec 2, 3 et 4 paramètres.**
- **Fonctions logistiques avec 5 paramètres, fonctions Gompertz : elles sont non symétriques.**
- **Fonctions logistique avec 2 paramètres : la réponse doit être entre 0 et 1.**

Courbes exponentielles de croissance et de décroissance

- **Fonctions exponentielles, double exponentielle (bi-exponentielle), fonction mécanique de croissance.**
- **Fonctions exponentielles avec 2 et 3 paramètres ; la fonction avec 3 paramètres a une asymptote.**
- **Fonctions bi-exponentielle possède 2 procédés de croissance ou de décroissance.**
- **Fonctions mécaniques et fonctions exponentielle avec 3 paramètres : ces fonctions sont toujours croissantes mais le taux de croissance diminue et les fonctions ont une asymptote.**

Courbes avec un pic

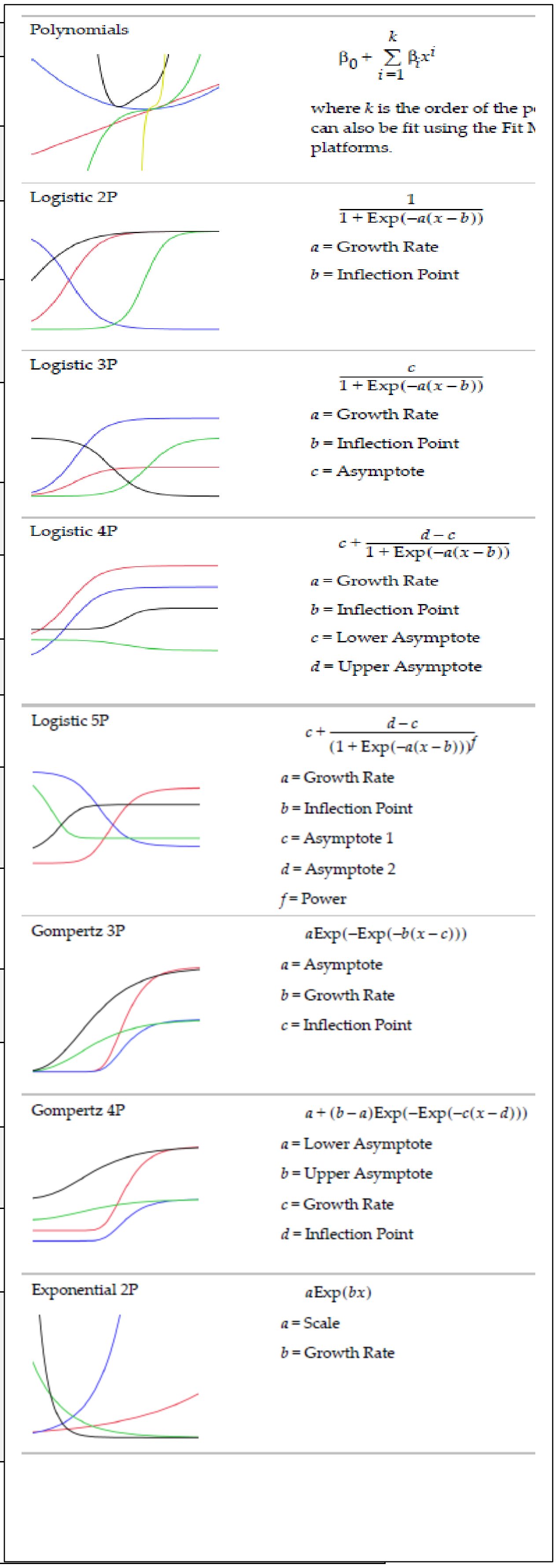
- **Fonctions gaussiennes, fonctions de Lorentz ; ces fonctions sont croissantes jusqu'à un pic et ensuite elles sont décroissantes.**
- **La fonction gaussienne est la densité de Gauss (normale) à un facteur d'échelle près.**
- **La fonction de Lorentz est la densité de Cauchy à un facteur d'échelle près.**

Courbes pharmacocinétique : étude de la concentration d'un médicament dans le corps

- **Fonctions de dose orale avec 1 compartiment ou 2 compartiments (Bolus).**
- **Fonctions bi-exponentielle avec 4 paramètres.**

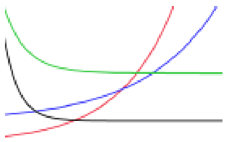
Courbes Michaelis-Menten en cinétique biochimique.

MODÈLE	FORMULE Y =	
Polynômes	$\beta_0 + \sum_{i=1}^{k} \beta_i x^i$	
Logistique 2P	$\frac{1}{1+\text{Exp}(-a(x-b))}$	a taux croissance b point d'inflexion
Logistique 3P	$\frac{c}{1+\text{Exp}(-a(x-b))}$	a taux croissance b point inflexion c asymptote
Logistique 4P	$c+\frac{d-c}{1+\text{Exp}(-a(x-b))}$	a taux croissance b point inflexion c asymptote inférieure d asymptote supérieure
Logistique 5P	$c+\frac{d-c}{(1+\text{Exp}(-a(x-b)))^f}$	a taux croissance b point inflexion c asymptote inférieure d asymptote supérieure f puissance
Gompertz 3P	$a\text{Exp}(-\text{Exp}(-b(x-c)))$	a asymptote b taux croissance c point Inflection
Gompertz 4P	$a+(b-a)\text{Exp}(-\text{Exp}(-c(x-d)))$	a asymptote inférieure b asymptote supérieure c taux croissance d pont inflection
Exponentielle 2P	aExp(bx)	a échelle b taux croissance
Exponentielle 3P	a + bExp(cx)	a asymptote b échelle c taux croissance
Bi-Exponentielle 4P	aExp(– bx) + cExp(– dx)	a échelle 1 b taux décroissance 1 c échelle 2 d taux décroissance 2
Bi-Exponentielle 5P	a + bExp(– cx) + dExp(– fx)	a asymptote b échelle 1 c taux décroissance 1 d échelle 2 f taux décroissance 2
Croissance mécanique	a(1 – bExp(– cx))	a asymptote b échelle c taux croissance
Sommet Gauss	$a\text{Exp}\left(-\frac{1}{2}\left(\frac{x-b}{c}\right)^2\right)$	a sommet b point critique c taux croissance
Sommet Lorentz (Cauchy)	$\frac{ab^2}{(x-c)^2+b^2}$	a sommet b taux croissance c point critique
dose orale 1 compartiment	$\frac{abc}{c-b}(\text{Exp}(-bx)-\text{Exp}(-cx$	a aire courbe b taux élimination c taux absorption
dose Bolus 2 compartiments	$\frac{a}{\alpha-\beta}((\alpha-b)\text{Exp}(-\alpha x)-(\beta-b)\text{Exp}(-\beta$ $\alpha=\frac{1}{2}(b+c+d+\sqrt{(b+c+d)^2-4bd})$ $\beta=\frac{1}{2}(b+c+d-\sqrt{(b+c+d)^2-4bd})$	a concentration initiale b taux transfert In c taux transfert out d taux élimination
Michaelis-Menten	$\frac{ax}{b+x}$	a taux max réaction b infinité adverse



Exponential 3P

$a + b\text{Exp}(cx)$



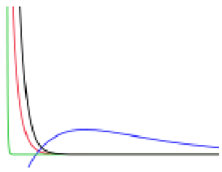
a = Asymptote

b = Scale

c = Growth Rate

Biexponential 4P

$a\text{Exp}(-bx) + c\text{Exp}(-dx)$



a = Scale 1

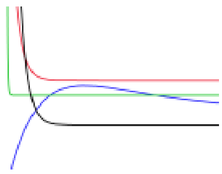
b = Decay Rate 1

c = Scale 2

d = Decay Rate 2

Biexponential 5P

$a + b\text{Exp}(-cx) + d\text{Exp}(-fx)$



a = Asymptote

b = Scale 1

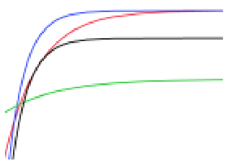
c = Decay Rate 1

d = Scale 2

f = Decay Rate 2

Mechanistic Growth

$a(1 - b\text{Exp}(-cx))$



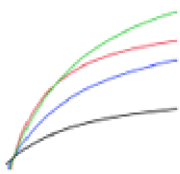
a = Asymptote

b = Scale

c = Growth Rate

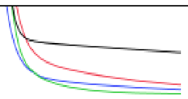
Michaelis-Menten

$\frac{ax}{b+x}$



a = Max Reaction Rate

b = Inverse Affinity



c = Transfer Rate Out

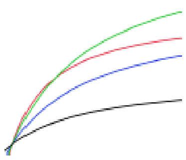
d = Elimination Rate

Michaelis-Menten

$\frac{ax}{b+x}$

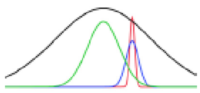
a = Max Reaction Rate

b = Inverse Affinity



Gaussian Peak

$a\text{Exp}\left(-\frac{1}{2}\left(\frac{x-b}{c}\right)^2\right)$



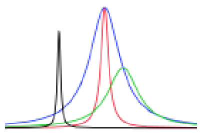
a = Peak Value

b = Critical Point

c = Growth Rate

Lorentzian Peak

$\frac{ab^2}{(x-c)^2 + b^2}$



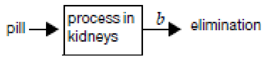
a = Peak Value

b = Growth Rate

c = Critical Point

One Compartment Oral Dose

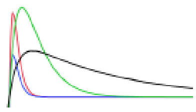
$\frac{abc}{c-b}(\text{Exp}(-bx) - \text{Exp}(-cx))$



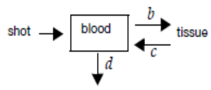
a = Area Under Curve

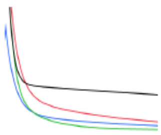
b = Elimination Rate

c = Absorption Rate



Model	Formula
Two Compartment IV Bolus Dose	$\frac{a}{\alpha-\beta}((\alpha-b)\text{Exp}(-\alpha x)-(\beta-b)\text{Exp}(-\beta x))$
	$\alpha = \frac{1}{2}(b+c+d+\sqrt{(b+c+d)^2 - 4bd})$
	$\beta = \frac{1}{2}(b+c+d-\sqrt{(b+c+d)^2 - 4bd})$





a = Initial Concentration

b = Transfer Rate In

c = Transfer Rate Out

d = Elimination Rate

Critères pour la comparaison de modèles :définition

AIC, AICc AICcWeigh, BIC, SSE, RMSE, R-square

(Source JMP17)

AIC critère d'information Akaike

AIC = - 2 log(L) + 2k = déviance + 2k

L : vraisemblance maximisée k : nombre de paramètres du modèle

La déviance du modèle (- 2 log(L)) est pénalisée par 2 fois le nombre de paramètres k. L'AIC représente un compromis entre le biais (qui diminue avec k) et la parcimonie (nécessité de décrire les données avec le plus petit nombre de paramètres possible).

AICc Akaike Information Criterion corrigé

AICc = AIC + [2 k(k - 1) / (n - k + 1)]

calcul avec SSE

sum of Square errors

$$\mathrm{AICc} = n\ln\left(\frac{SSE}{n}\right) + 2k + 2k(k+1)/(n-k-1) + n\ln(2\pi) + n$$

AICc est une modification du critère Akaide pour des petits échantillons (n < 40k). On doit avoir n ≥= k + 2. AICc est préférable à AIC pour de petits échantillons. Le critère est une mesure de l'ajustement d'un modèle statistique. AICc permet de comparer 2 modèles ou plus. Le modèle ayant la plus petite valeur AICc est le meilleur.

AICcWeight Normalisation des valeurs AICc pour avoir une somme de 1

Peut s'interpréter comme la probabilité que le modèle est le meilleur parmi les modèles ajustés. Le modèle ayant une valeur AICc weight la plus près de 1 est le modèle ayant le meilleur ajustement.

BIC Bayesian Information Criterion

BIC = -2 * Log(L) + k * log(n)

mesure de comparaison entre des modèles. Le modèle ayant la plus petite valeur de BIC est le meilleur.

SSE : Sum of Square of Errors

Somme des carrés des différences entre les valeurs observées et les valeurs prédites.

MSE Mean Square Error valeur moyenne de SSE.

RMSE : Root Mean Square Error Racine carrée de MSE.

R-Square

Proportion de la variation de la variable de réponse expliquée par le modèle.
Le modèle ayant la valeur la plus près de 1 est le meilleur.

Choix du critère Meilleur critère

Les critères AIC, AICc, AiCcWeigh sont préférés à R-square car ils tiennent en compte le nombre de paramètres.

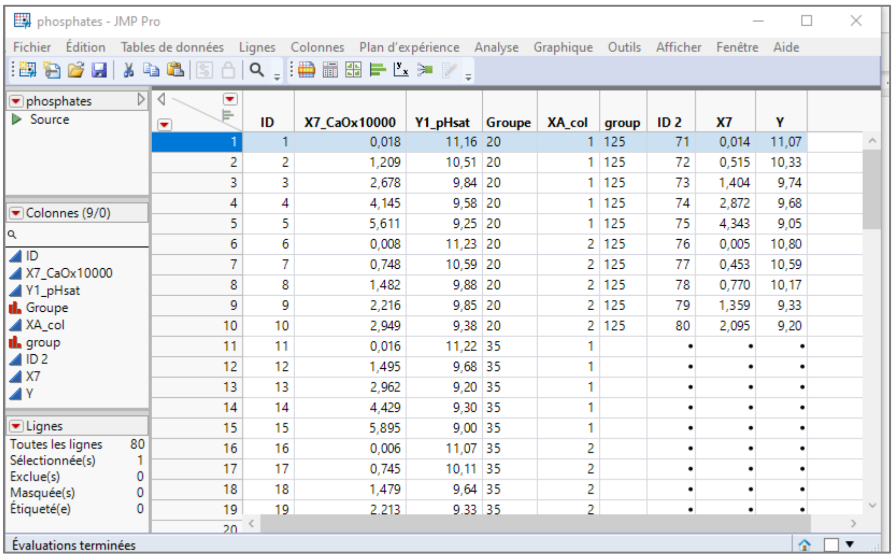
Exemple avec JMP

Data : Phosphates.jmp

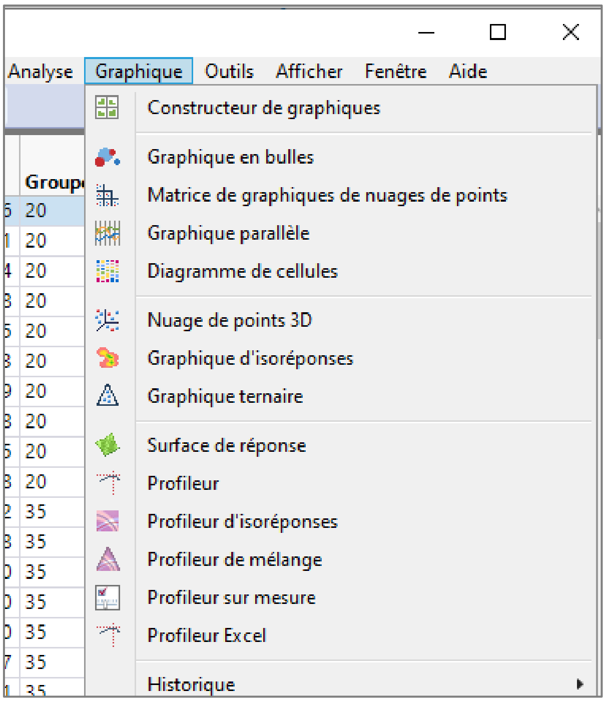
Group	ID2	X7	Y
125	71	0,014	11,07
125	72	0,515	10,33
125	73	1,404	9,74
125	74	2,872	9,68
125	75	4,343	9,05
125	76	0,005	10,80
125	77	0,453	10,59
125	78	0,770	10,17
125	79	1,359	9,33
125	80	2,095	9,20

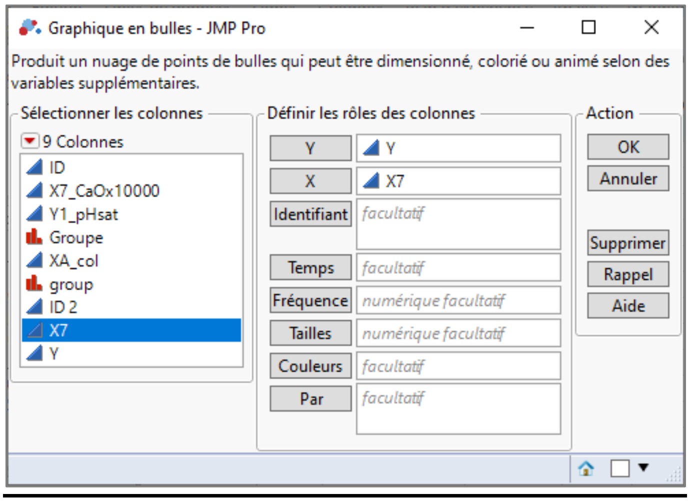
Exécution avec JMP : étapes

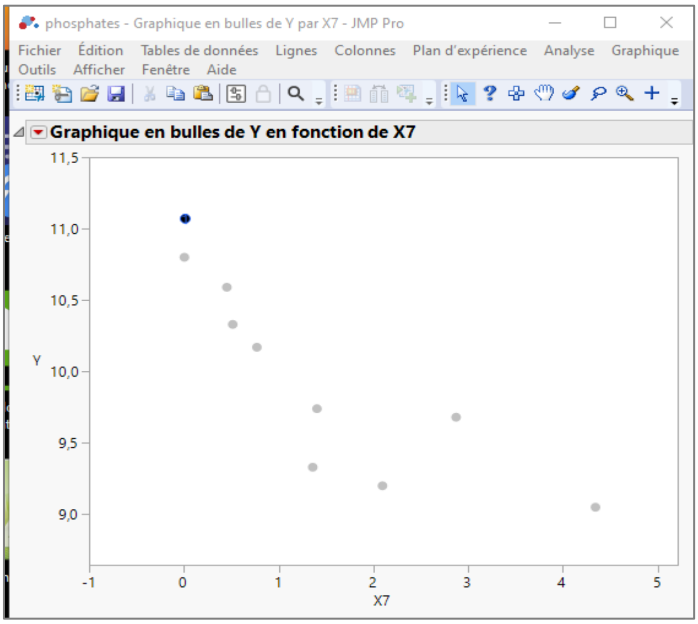
Étape1 : création d'un fichier phosphates.jmp



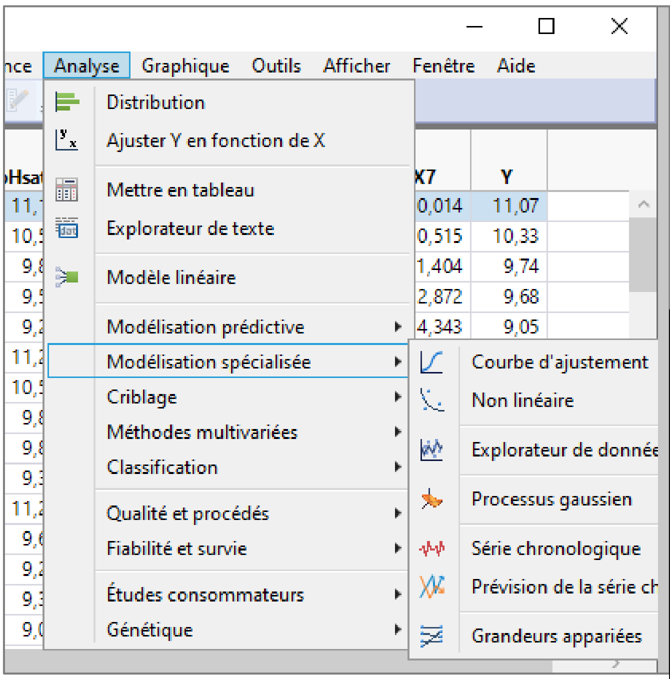
Étape2 : visualisation des données

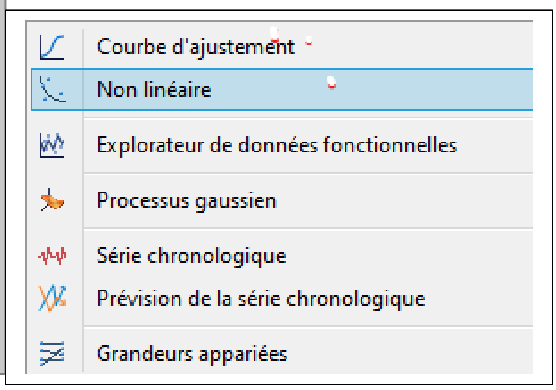


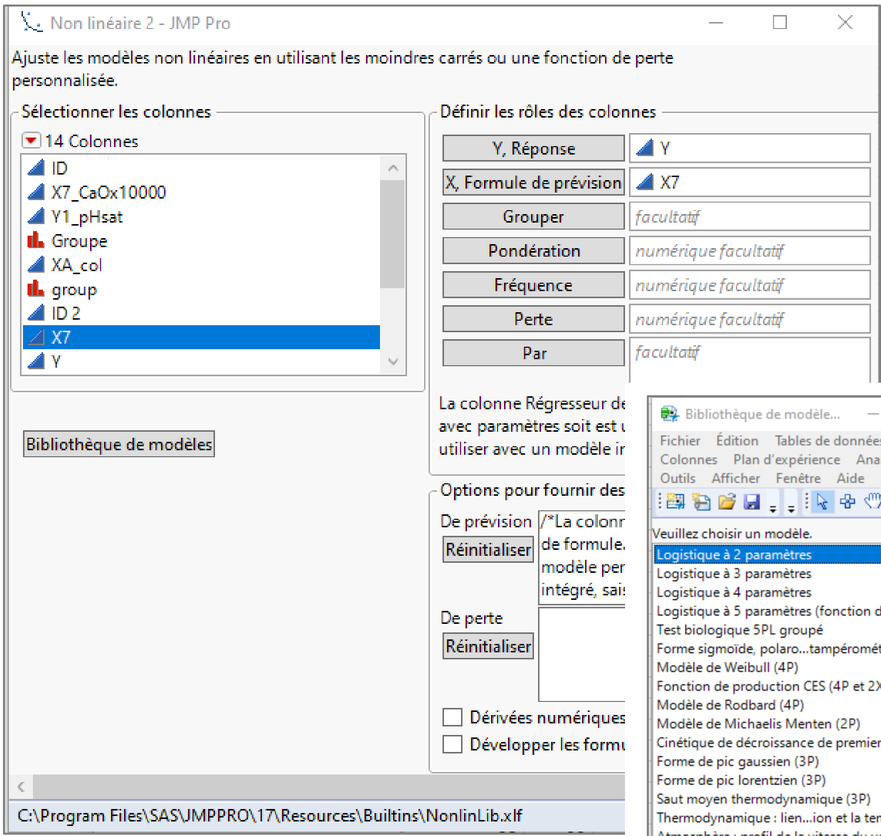


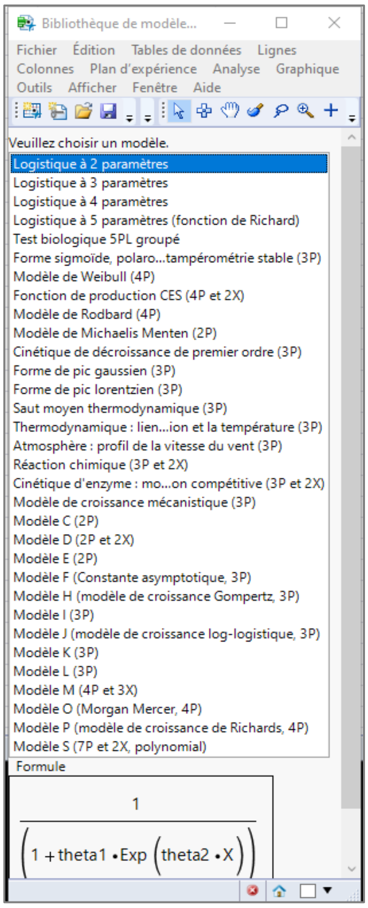


Étape3 : modélisation non linéaire









La version 17 de JMP possède une liste de 33 modèles préprogrammés en cliquant sur l'onglet *Bibliothèque de modèles*

Il est aussi possible de définir une fonction de son choix.

Voici le tableau de la comparaison de 13 modèles choisis et ajustées aux données.

Comparaison de modèles

Modèle	AICc	Poids d'Akaike corrigé (AICc)	BIC	Somme des carrés des écarts (SSE)	Erreur quadratique moyenne (MSE)	RMSE	R carré
Exponentielle 2 paramètres	0,1834632	0,2580393	-2,908781	0,2193809	0,0274226	0,1655978	0,9490462
Exponentielle 3 paramètres	0,5880499	0,2107808	-6,20161	0,1253699	0,01791	0,1338282	0,9708814
Croissance mécanistique	0,5880499	0,2107808	-6,20161	0,1253699	0,01791	0,1338282	0,9708814
Quadratique	1,1286939	0,1608539	-5,660966	0,1323345	0,0189049	0,1374952	0,9692638
Linéaire	1,6480929	0,1240639	-1,444152	0,2539843	0,031748	0,1781798	0,9410092
Logistique à 3 paramètres	6,1879017	0,0128185	-0,601758	0,2194783	0,031354	0,1770707	0,9490236
Gompertz à 3 paramètres	6,3363229	0,0119017	-0,453337	0,2227601	0,0318229	0,1783897	0,9482614
Cubique	8,7504938	0,0035594	-4,736581	0,1152971	0,0192162	0,1386225	0,9732209
Biexponentielle à 4 paramètres	9,4463254	0,0025135	-4,040749	0,1236056	0,0206009	0,1435303	0,9712912
Logistique à 4 paramètres	9,5880503	0,0023416	-3,899024	0,1253699	0,020895	0,144551	0,9708814
Gompertz à 4 paramètres	9,5880539	0,0023416	-3,899021	0,1253699	0,020895	0,144551	0,9708814
Biexponentielle à 5 paramètres	22,545817	3,5955e-6	-3,638672	0,1022116	0,0204423	0,1429766	0,9762602
Logistique à 5 paramètres	24,357388	1,4534e-6	-1,827101	0,1225112	0,0245022	0,1565319	0,9715454

Ce tableau de comparaison provenant de la version 16 de JMP ne semble pas disponible dans la version 17.