

Modèles empiriques et modèles mécanistes

et divers outils de modélisation

Marc Girondot, Université Paris Saclay

marc.girondot@universite-paris-saclay.fr



1

Plan

- Modèle empirique
 - Définition, avantage, inconvénient
- Modèle mécaniste
 - Définition, avantage, inconvénient
- Un exemple de modèle
 - La croissance d'un embryon sous contrainte de l'environnement
 - Modèle empirique
 - Modèle mécaniste

2

Modèle empirique

- Une loi empirique est une loi ou une formule issue de faits expérimentaux, ou validée par l'expérience, mais dont on ne connaît pas de base théorique, ou qu'on ne peut pas relier à une base théorique simple.
- Un modèle empirique est donc l'utilisation d'une loi empirique c'est-à-dire reflétant directement les observations.

3

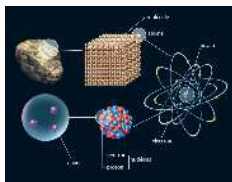
Modèle empirique

- Avantages
 - Même si on ne comprend pas le fonctionnement du système, il est quand même possible de faire des prédictions ou des inférences;
- Inconvénients
 - Si on ne reste pas exactement dans la gamme des paramètres qui ont permis de générer les données, il est impossible de vérifier si le comportement du modèle est correct.

4

Modèle mécaniste

- Dans un modèle mécaniste, on essaye de modéliser le fonctionnement d'un système. Le niveau d'organisation choisit dépend à la fois des données disponibles et du niveau auquel on attend des réponses. On prend en plus en compte le temps de calcul nécessaire pour exécuter le modèle.



5

Modèle mécaniste

- Avantages
 - On est au plus proche du fonctionnement du système; on pourra explorer des gammes de paramètres pour lesquelles on n'a pas d'information.
- Inconvénients
 - Le nombre de paramètres peut être très important. Par exemple, le modèle météorologique mécaniste de l'ECMWF utilise 5300 paramètres et le temps de calcul pour calculer le climat à la surface de la terre (résolution $0,75^\circ \times 0,75^\circ$) se compte en heure.

ECMWF is the European Centre for Medium-Range Weather Forecasts
<https://www.ecmwf.int>

6

Un modèle de croissance de l'embryon sous contrainte de l'environnement

- Exploration des données
- Modèle empirique
- Modèle mécaniste

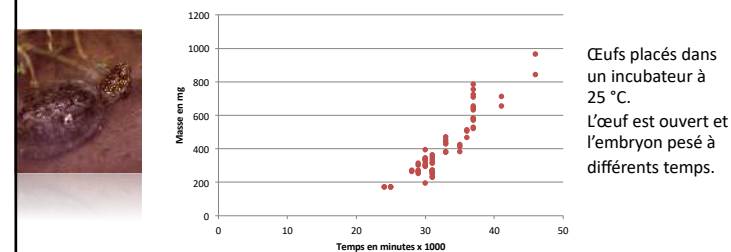
Notions qui seront vues:

- Maximum de vraisemblance
- AIC, Akaike weight, sélection de modèle, AICc, BIC
- Equation différentielle
- Approximation par la méthode de Runge-Kutta

7

Croissance de l'embryon

- Pensez à observer et interpréter les observations

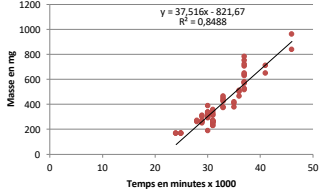


Comment modéliser ces données ?

8

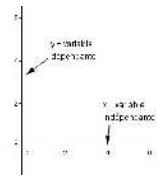
Croissance de l'embryon

- Pensez à observer et interpréter les observations



Quand on fait une régression linéaire au sens des moindres carrés, on fait des hypothèses:

- Qu'il existe une relation linéaire entre x et y ;
- Que la variable en x ne soit pas une variable aléatoire;
- Que la distribution marginale en y soit normale;
- Que la variance de la distribution marginale en y soit constante.



9

- Dans l'ajustement par les moindres carrés, on cherche les paramètres a et b qui minimisent la somme des carrés entre les observations et les prédictions:

$$y_{pred} = a x + b$$

$$SSQ = \sum (y_{obs} - y_{pred})^2$$

Pourquoi des carrés ?

Le premier à avoir utilisé une méthode similaire est Ruder Josip Bošković, (18 mai 1711, Raguse [en croate, Dubrovnik] – 13 février 1787, Milan) qui utilisait des valeurs absolues.

10

Probabilité, fréquence, vraisemblance

INTERLUDE


- Il existe de nombreuses définitions de ces termes qui ne recouvrent donc pas exactement la même chose chez différents auteurs.
 - En général, le terme probabilité est utilisé pour décrire la chance (ou le risque !) qu'un événement particulier se produise.
 - Le terme fréquence (ou plus précisément fréquence empirique) désigne une réalisation particulière de la probabilité qu'un événement particulier se produise.
 - Le terme vraisemblance désigne une probabilité conditionnelle, c'est à dire la probabilité d'observations en fonction de paramètres supposés connus.
- En anglais, *frequency* est utilisé pour nommer des dénombrements (par exemple sur la commande hist()) de R, *freq=TRUE* permet d'afficher les nombres d'observations), on distingue:
 - *Frequency (or absolut frequency)*, des nombres
 - *Relative frequency (or empirical probability)*, des fréquences !

11

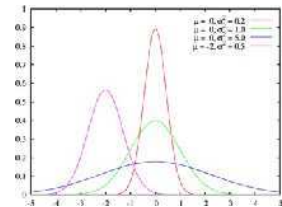
Pierre-Simon de Laplace, Essai sur la philosophie des probabilités, 1820

Loi normale ou de Gauss-Laplace

Laplace (1749-1827) et Gauss (1777-1855)



- Soit une loi normale $\mathcal{N}(\mu, \sigma)$
- La vraisemblance d'une observation Y_i tirée dans cette loi normale est



$$L\{Y_i; m\} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(Y_i - \mu)^2}{2\sigma^2}\right)$$

$-\ln L\{Y; m\}$

Pourquoi travailler en \ln ?

12

logarithmes

$\log_a(x y) = \log_a(x) + \log_a(y)$
 $\log_a(a) = 1 \quad \log_a(1) = 0$
 $\log_a(x / y) = \log_a(x) - \log_a(y)$
 $\log_a(x^y) = y \log_a(x)$

p	0	1
$\ln p$	-∞	0
$-\ln p$	+∞	0

« Maximum de vraisemblance »

Ces règles s'appliquent quelque soit la base a choisie pour le logarithme.

13

Approximations

$$S_1 = \frac{2^{(K_1-1)} K_1 S}{2^{K_1} - 1}$$

- Comment calculer S_1 si K_1 est très grand ?

14

$$S_1 = \frac{2^{(K_1-1)} K_1 S}{2^{K_1} - 1}$$

- Si K_1 est très grand, 2^{K_1} dépasse les capacités de calcul d'un ordinateur; on passe en log:
- $\log(2^{(K_1-1)} K_1 S) - \log(2^{K_1} - 1)$
- $(K_1 - 1) \log(2) + \log(K_1 S) - \log(2^{K_1})$
- $(K_1 - 1) \log(2) + \log(K_1 S) - K_1 \log(2)$
- $\log(2)(K_1 - 1 - K_1) + \log(K_1 S) = \log(K_1 S) - \log(2)$
- $\log\left(\frac{K_1 S}{2}\right)$ d'où en prenant l'exponentielle: $S_1 = \frac{K_1 S}{2}$


> K1 = 267.7304; S = 2.509569 > (2^(K1 - 1)*K1*S)/(2^K1 - 1) [1] 335.9439 > K1*S/2 [1] 335.9439	> K1=1200; S=3 > (2^(K1 - 1)*K1*S)/(2^K1 - 1) [1] NaN > K1*S/2 [1] 1800
---	---

15


Série de Taylor

Soit f une fonction d'une variable réelle ou complexe, indéfiniment dérivable en un point a . La série de Taylor de f en ce point est la série de fonctions :

$$f(a) + \frac{f'(a)}{1!}(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \frac{f'''(a)}{3!}(x - a)^3 + \dots$$



Brook Taylor, 1685-1731



Colin Maclaurin, 1698 - 1746

Si $a = 0$, la série est aussi appelée la série de Maclaurin de f .

$$\log(1 + x) = x - x^2/2 + x^3/3 - x^4/4 \dots$$


$$\log(1 - x) = -x - x^2/2 - x^3/3 - x^4/4 \dots$$

16

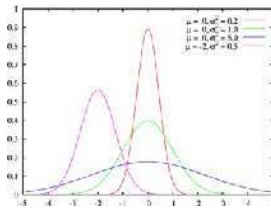
Pierre-Simon de Laplace, Essai sur la philosophie des probabilités, 1820

Loi normale ou de Gauss-Laplace

Laplace (1749-1827) et Gauss (1777-1855)



- Soit une loi normale $\mathcal{N}(\mu, \sigma)$
- La vraisemblance d'une observation Y_i tirée dans cette loi normale est



$$L\{Y_i; m\} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(Y_i - \mu)^2}{2\sigma^2}\right)$$

$$-\ln L\{Y_i; m\} = \ln(\sigma) + \frac{1}{2}\ln(2\pi) + \frac{(Y_i - \mu)^2}{2\sigma^2}$$

17

Soit un échantillon de k points

Si σ est constant

$$-\ln L\{Y; m\} = k \left(\ln(\sigma) + \frac{1}{2}\ln(2\pi) \right) + \frac{1}{2\sigma^2} \sum_{i=1}^k (Y_i - \mu)^2$$

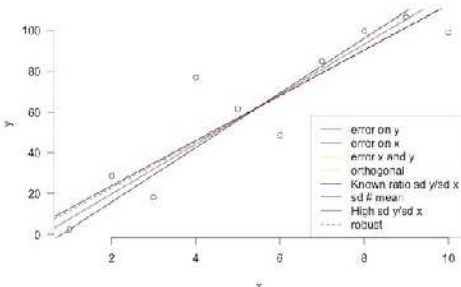
Notez que le premier terme est une constante qui ne dépend pas des observations.
Le second terme est une somme des moindres carrés pondérés par la variance qui est elle-même une constante.

Donc la régression au sens des moindres carrés recherche les paramètres du modèle (ici a et b) qui maximisent la vraisemblance des observations dans le modèle sous les hypothèses suivantes :

- Qu'il existe une relation linéaire entre x et y ;
- Que la variable en x ne soit pas une variable aléatoire;
- Que la distribution marginale en y soit normale;
- Que la variance de la distribution marginale en y soit constante.

18

Les régressions linéaires



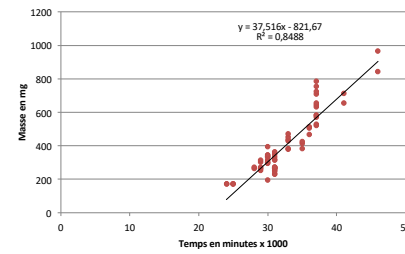
<https://biostatsr.blogspot.com/2018/10/the-mess-of-linear-regressions.html>

* Mess = pagaille

19

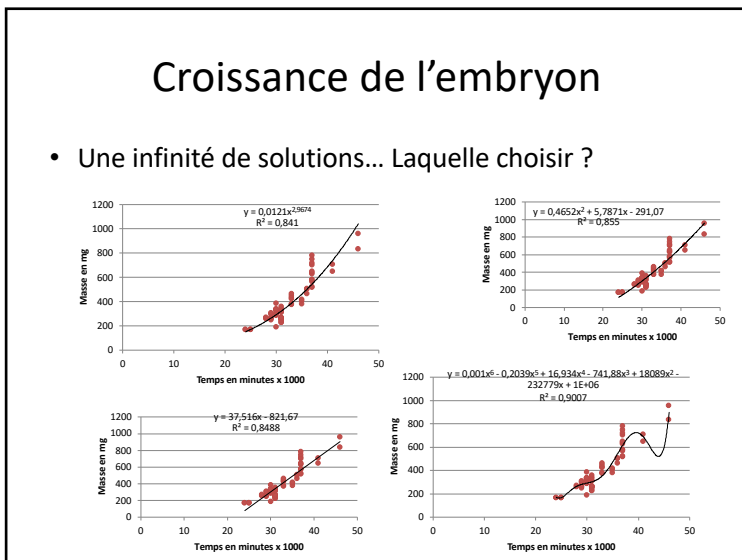
Croissance de l'embryon

- Pensez à observer et interpréter les observations

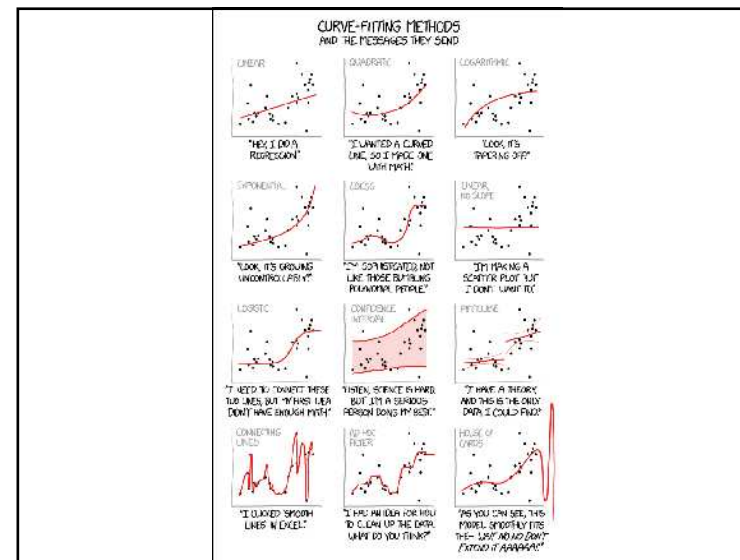


Régression linéaire ? Des masses négatives sont attendues à la ponte !
Problème classique de l'extrapolation

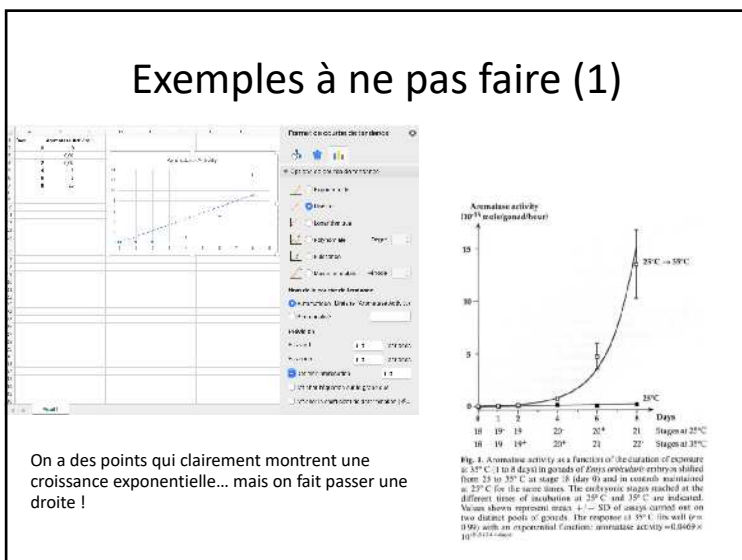
20



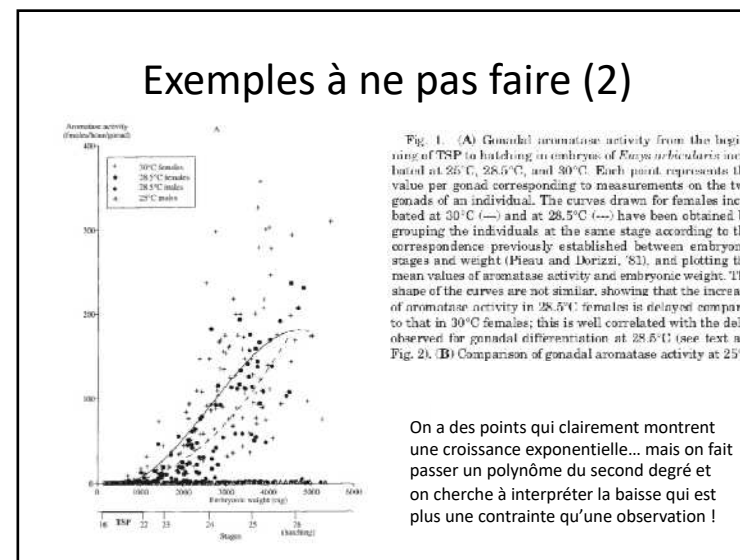
21



22



23



24

Qualité d'un ajustement

- L'ajustement d'une courbe théorique sur un ensemble d'observations se fait selon un critère d'optimisation, par exemple la somme des carrés de la différence entre les observations et les prédictions (« *méthode des moindres carrés* »).
- **Mais comment savoir qu'un ajustement est meilleur qu'un autre ?**

25

Qualité d'un ajustement

- De nombreuses études utilisent la moyenne du carré des écarts (MSE) et sa racine carrée (RMSE), ou l'erreur absolue moyenne (MAE) et sa variante en pourcentage (MAPE).
 - Bien qu'utiles, une seule valeur d'entre eux ne dit pas grand-chose sur les performances de la régression par rapport à la distribution originale.

Chicco, D., Warrens, M.J., Jurman, G., 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Comput Sci 7, e623.

26

Corrélation

- En probabilités et en statistique, la corrélation entre plusieurs variables aléatoires ou statistiques est une notion de liaison qui mesure leur indépendance.
- La corrélation linéaire entre variables quantitatives mesure l'ajustement d'une variable par rapport à l'autre par une relation affine obtenue par régression linéaire. Pour cela, on calcule un coefficient de corrélation linéaire, quotient de leur covariance par le produit de leurs écarts types.

$$r = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}}$$

27

Corrélation

- Fisher (1915) et plus tard Olkin and Pratt (1958) ont fourni des estimateurs approximativement sans biais de la corrélation de population (en prenant en compte que des échantillons aléatoires normaux ont une corrélation non nulle).

$$\hat{\rho} = r \left(1 + \frac{1 - r^2}{2n} \right)$$

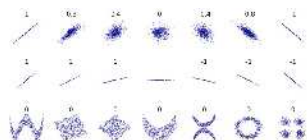
$$\hat{\rho} = r \left(1 + \frac{1 - r^2}{2(n-3)} \right)$$

Fisher, R.A., 1915. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. Biometrika 10, 507-521.
Olkin, I., Pratt, J.W., 1958. Unbiased estimation of certain correlation coefficients. The Annals of Mathematical Statistics 29, 201-211.

28

Corrélation

- La corrélation ne mesure que la relation linéaire entre deux variables, pas la relation elle-même:



29

Qualité d'un ajustement

- On peut aussi utiliser le coefficient de détermination, appelé R^2 .

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \in [0, 1] \quad R^2 = \text{corr}(y, \hat{y})^2$$

- Le coefficient de détermination représente la proportion de la variance des données observées (y_i) expliquée par le modèle (cas d'un modèle linéaire avec une distribution normale et une variable) produisant des prédictions \hat{y}_i .
- Mais la valeur du critère d'ajustement est dépendante du nombre de paramètres dans le modèle. Si on a autant de paramètres que d'observations, l'ajustement sera parfait:
 - R^2 vaudra 1
 - La somme des carrés des écarts au modèle vaudra 0

30

Le R^2 ajusté

- Ezekiel (1929) a proposé une correction du R^2 pour prendre en compte le nombre de paramètres p et la quantité de données n :

$$R_{adj}^2 = 1 - (1 - R^2) \frac{n - 1}{n - p - 1}$$

- Le R^2 ajusté peut être négatif, et sa valeur sera toujours inférieure ou égale à celle de R^2 .
- Mais attention, le R^2 ajusté ne sert pas à mesurer la force d'une relation mais à classer des modèles.

Ezekiel, M., 1929. The application of the theory of error to multiple and curvilinear correlation. Journal of the American Statistical Association 24, 99-104.

31

Généralisation du R^2

- Une généralisation du R^2 a été proposée par Nagelkerke (1991), avec $L(0)$ la vraisemblance du modèle nul et $L(\hat{\beta})$ la vraisemblance du modèle ajusté:

$$R^2 = 1 - (L(0)/L(\hat{\beta}))^{2/n}$$

- Il est cohérent avec le R^2 classique, c'est-à-dire la définition générale appliquée (par ex. la régression linéaire) donne le R^2 classique.
- Il est cohérent avec le maximum de vraisemblance comme méthode d'estimation, c'est-à-dire que les estimations du maximum de vraisemblance des paramètres du modèle maximisent R^2 .
- Il est asymptotiquement indépendant de la taille de l'échantillon n .
- $1 - R^2$ a l'interprétation de la proportion de « variation » inexpliquée.
- Il est sans dimension, c'est-à-dire qu'il ne dépend pas des unités utilisées.
- Dans le cas d'un modèle binomial, $\max(R^2) = 1 - L(0)^{2/n}$; on utilisera donc plutôt $R^2 = R^2 / \max(R^2)$

Nagelkerke, N.J.D., 1991. A note on a general definition of the coefficient of determination. Biometrika 78, 691-192.

32

Le rasoir d'Ockam



- Le rasoir d'Ockham est un principe de raisonnement philosophique entrant dans les concepts de rationalisme. Son nom vient du philosophe franciscain Guillaume d'Ockham (1285, Ockham, Royaume-Uni - 1347, Munich, Allemagne).
- On le trouve également appelé principe de simplicité, principe d'économie ou principe de parcimonie (en latin *lex parsimoniae*).
« **Les multiples ne doivent pas être utilisés sans nécessité.** »

Il reprend un adage scolastique dérivé d'Aristote : « **C'est en vain que l'on fait avec plusieurs ce que l'on peut faire avec un petit nombre** » qui d'ailleurs est cité par Guillaume d'Ockham (*Summa totius logicae*, I, 12) (1323).

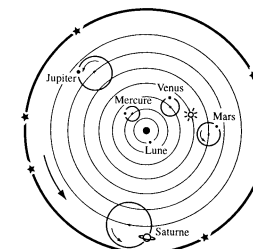
33

Claude Ptolémée, né vers 100 - mort vers 168 à Canope (Egypte).

Modèle géocentrique



- Le mouvement des planètes est expliqué par Ptolémée par une série de mouvements circulaires avec la terre au centre du système, soit un modèle géocentrique.



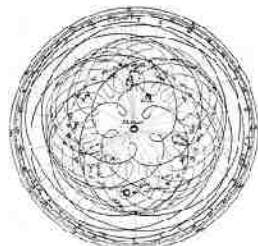
34

Claude Ptolémée, né vers 100 - mort vers 168 à Canope (Egypte).

Modèle géocentrique



- Pour expliquer les connaissances acquises au cours des siècles suivants, le modèle nécessitait des séries importantes d'orbites autour des orbites d'ordre inférieur.



35

Les sphères armillaires

- Ce modèle est utilisé pendant très longtemps et encore au 16^{ème} siècle.



Jan Gossaert (ca 1478 – 1532), Portrait de jeune princesse portant une sphère armillaire, vers 1530, The National Gallery, Londres



36



37


Les méthodes de sélection de modèles

- Une *fonction de vraisemblance* (souvent simplement appelée la *vraisemblance*) est une fonction d'un modèle statistique utilisant des paramètres le tout dépendant de données observées.


39

Galilée, né à Pise en 1564 et mort à Arcetri près de Florence le 8 janvier 1642 (77 ans).

Modèle héliocentrique



- On va donc chercher utiliser un critère de meilleur ajustement mais en pénalisant par le nombre de paramètres utilisés dans le modèle pour arriver à la proposition de Copernic, reprise par Galilée, du modèle héliocentrique.



Nicolas Copernic est un astronome polonais, également chanoine, médecin et mathématicien, né le 19 février 1473 à Thorn (Toruń), Prusse royale (royaume de Pologne) et mort le 24 mai 1543 à Frauenburg Prusse royale (royaume de Pologne).

38

Vraisemblance

- La fonction de vraisemblance, notée $L(x_1, \dots, x_n \mid \theta_1, \dots, \theta_k)$ est une fonction de probabilités conditionnelles qui décrit les valeurs x_i (observations) d'une loi statistique en fonction des paramètres θ_j supposés connus (mais potentiellement modifiables dans le cas d'un ajustement par le maximum de vraisemblance).
- Elle s'exprime à partir de la fonction de densité $f(x \mid \theta)$ par

$$L(x_1, x_2, \dots \mid \theta) = \prod_{i=1}^n f(x_i; \theta)$$

40

Maximum de vraisemblance

- Utilisation du maximum de vraisemblance pour déterminer les paramètres d'une distribution.
- Etant donné un échantillon observé x et une fonction de vraisemblance f , la vraisemblance quantifie la probabilité que les observations proviennent effectivement d'un échantillon (théorique) de la loi f .
- Estimer un paramètre par la méthode du maximum de vraisemblance, c'est proposer comme valeur de ce paramètre celle qui rend maximale la vraisemblance, à savoir la probabilité d'observer les données comme réalisation d'un échantillon de la loi f .

41

Vraisemblance des modèles

Modèle	a+bx	a*x^b	a+bx+cx^2	a+bx+cx^2+dx^3+ex^4+fx^5
-Ln L	323,21	324,13	322,20	320,20

En terme de vraisemblance, le modèle qui a la meilleure vraisemblance (-ln L plus faible) est le polynôme du cinquième degré.

Est-ce étonnant ? $L\{Y_i; m\} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(Y_i - \mu)^2}{2\sigma^2}\right)$



Pierre-Simon de Laplace, Essai sur la philosophie des probabilités, 1820

42

Vraisemblance des modèles

Modèle	a+bx	a*x^b	a+bx+cx^2	a+bx+cx^2+dx^3+ex^4+fx^5
-Ln L	323,21	324,13	322,20	320,20

En terme de vraisemblance, le modèle qui a la meilleure vraisemblance (-ln L plus faible) est le polynôme du cinquième degré.

Est-ce étonnant ?

$$L\{Y_i; m\} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(Y_i - \mu)^2}{2\sigma^2}\right)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \in [0, 1]$$

Ce n'est pas étonnant car le critère de minimisation utilisé dans la fonction de vraisemblance est proportionnel au R^2 . Il souffre du même problème que décrit précédemment: sensibilité au nombre de paramètres dans le modèle.

43

Comparer des modèles

Soit m_1 et m_2 , deux modèles avec p_1 et p_2 paramètres ajustés sur un jeu de données d et ayant pour vraisemblance L_1 et L_2 , on va définir la statistique AIC comme étant:

$$AIC = -2 \ln L + 2 p$$

Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19:716-723.

Entre deux modèles, on choisira donc l'AIC (Akaike Information Criterion) est le plus faible: meilleur ajustement mais pas trop de paramètres.

44

Comparer des modèles

L'AIC n'est pas un test au sens de la théorie des tests.
On peut convertir l'AIC en une statistique qui représente la probabilité que ce soit réellement le meilleur modèle parmi ceux testés.
C'est l'Akaike weight:

$$\frac{\exp\left(-\frac{1}{2}(AIC_i - AIC_{min})\right)}{\sum_{j=1}^k \exp\left(-\frac{1}{2}(AIC_j - AIC_{min})\right)}$$

Burnham KP, Anderson DR (2002) Model selection and multimodel inference: A practical information-theoretic approach. Springer-Verlag, New York

45

AIC, AICc, BIC

- Il existe d'autres critères que l'AIC.
 - L'AICc est une version corrigée de l'AIC pour les petits effectifs. Il évite une surparamétrisation.
 - Le BIC (Bayesian Information Criterion) peut-être utilisé lorsque le « vrai » modèle fait partie des modèles testés.


There is no true model.
Anderson & Burnham 1999

46

BIC

- Dans des cas particuliers, on peut se trouver en présence du « vrai » modèle. Par exemple si le test est du genre:
 - 'C'est différent' ou 'c'est identique'.
 - Le « vrai » modèle est alors présent puisque les deux modalités sont les seules possibles.

Girondot, M. & Guillon, J.-M. (2018) The w-value: An alternative to t- and X² tests. *Journal of Biostatistics & Biometrics*, 1, 1-4.



47

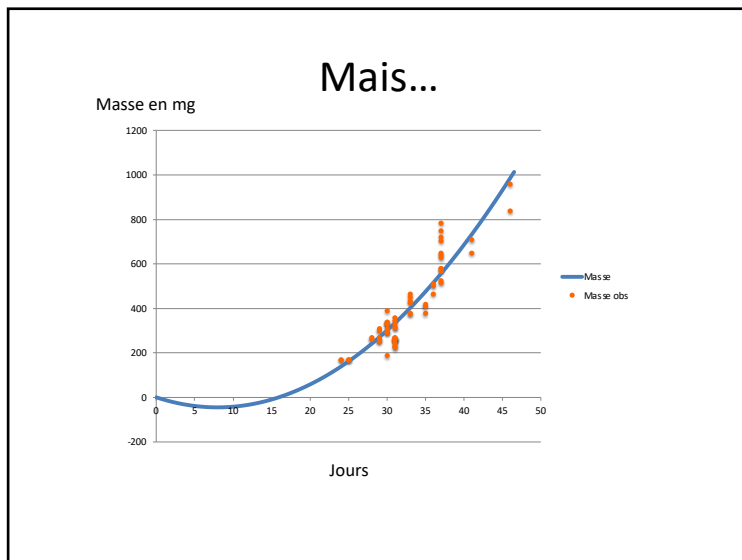
Vraisemblance des modèles

Modèle	a+bx	a*x^b	a+bx+cx^2	a+bx+cx^2+dx^3+ex^4+fx^5
-ln L	323,21	324,12	322,19	320,19
AIC	652,42	654,25	652,39	654,39
Akaike weight	0,359	0,143	0,363	0,133

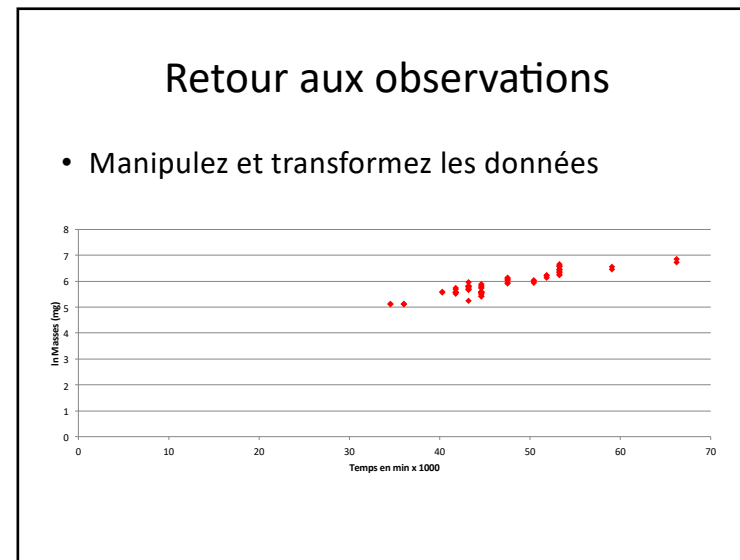
En terme d'AIC, le meilleur modèle est le polynôme du second degré. Il y a 36,3% de chance que ce soit réellement le meilleur modèle parmi ceux-testés.
Mais la droite a aussi 35,9% de chance d'être le meilleur modèle !

Attention, c'est une sélection de modèle, ce n'est pas un test.

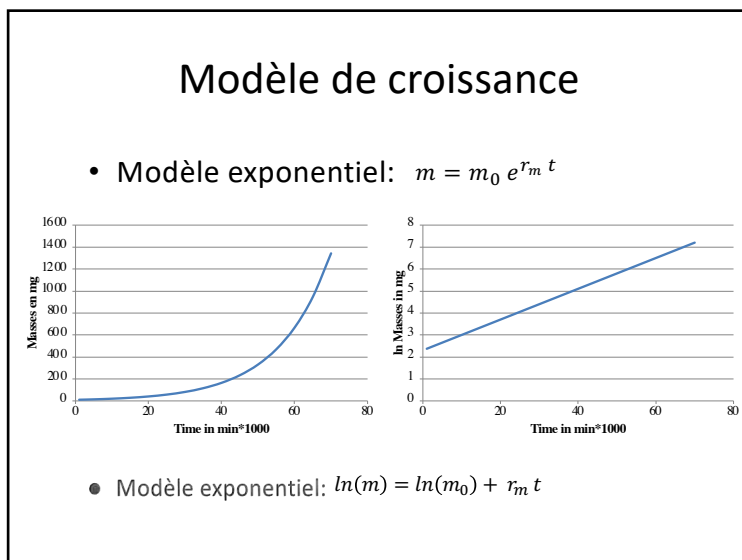
48



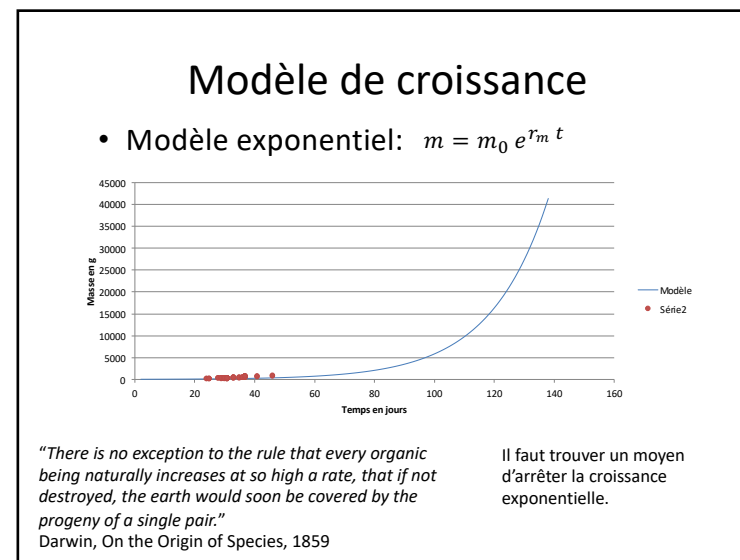
49



50






51



52

Différentes équations

- Verhulst P.F. 1838. Notice sur la loi que la population suit dans son accroissement. $y = a / (1 + \exp(b(c-t)))$
Correspondance Mathématique et Physique 10:113-121.
- Bertalanffy, L. von 1938. A quantitative theory of organic growth. *Hum. Biol.*, 10 (2): 181-213. $y = A [1 - B \exp(-kt)]^{1/(1-m)}$
- Gompertz B. 1825. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philos Trans R Soc Lond [Biol]* 115:513-585. $x = Ke^{be^{-at}}$
 $b = \ln(\frac{x_0}{K})$

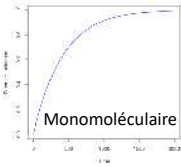
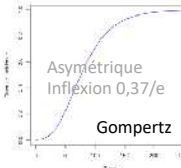
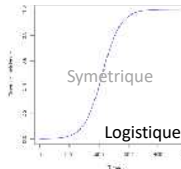
53

Généralisation

Richards F.J. 1959. A flexible growth function for empirical use. *Journal of Experimental Botany* 29:290-300.

$$g = a(1 + b \exp(-rt))^{1/(1-k)}$$

$$g = \begin{cases} a(1 + be^{-rt}) & \text{Monomolecular curve when } k = 0 \\ ae^{-be^{-rt}} & \text{Gompertz's curve when } k = -1 \\ \frac{a}{1 + be^{-rt}} & \text{Autocatalytic or logistic curve when } k = -2. \end{cases}$$

Façon astucieuse de tester différentes formulations de modèles qu'on peut changer avec un paramètre.

54

Fonction de Gompertz

- La fonction de Gompertz a une dynamique facilement interprétable par rapport aux valeurs de paramètres

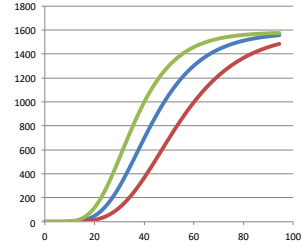
$$x = Ke^{be^{-at}} \quad b = \ln\left(\frac{x_0}{K}\right)$$

55

Fonction de Gompertz

- x_0 est la valeur à $t=0$
- K est l'asymptote

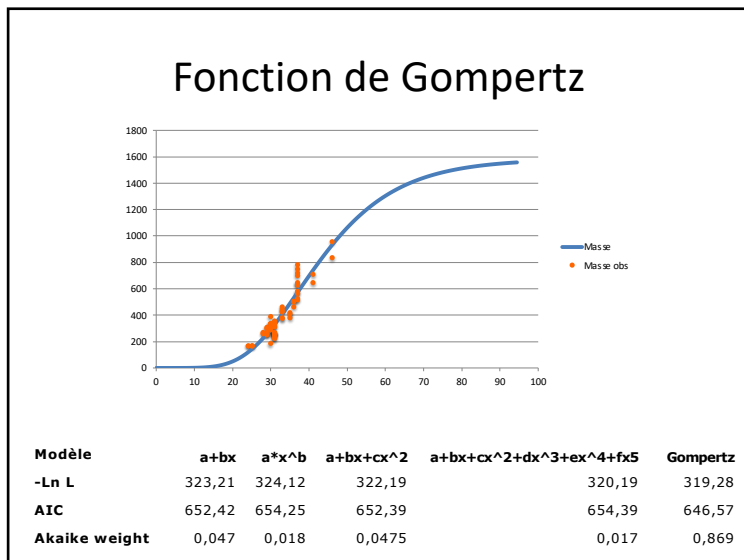
$$x = Ke^{be^{-at}} \quad b = \ln\left(\frac{x_0}{K}\right)$$



a est une mesure de la vitesse de la croissance

- $a=0,05$
- $a=0,04$
- $a=0,06$

56



57

Modèle de Gompertz

d'une description empirique à un modèle mécaniste

La taille d'une population est fonction du nombre d'individus :

- Verhulst, P.F., 1838. Notice sur la loi que la population suit dans son accroissement. Correspondance Mathématique et Physique 10, 113-121.
- Berec, L., 2002. Techniques of spatially explicit individual-based models: construction, simulation, and mean-field analysis. Ecological Modelling 150, 55-81.

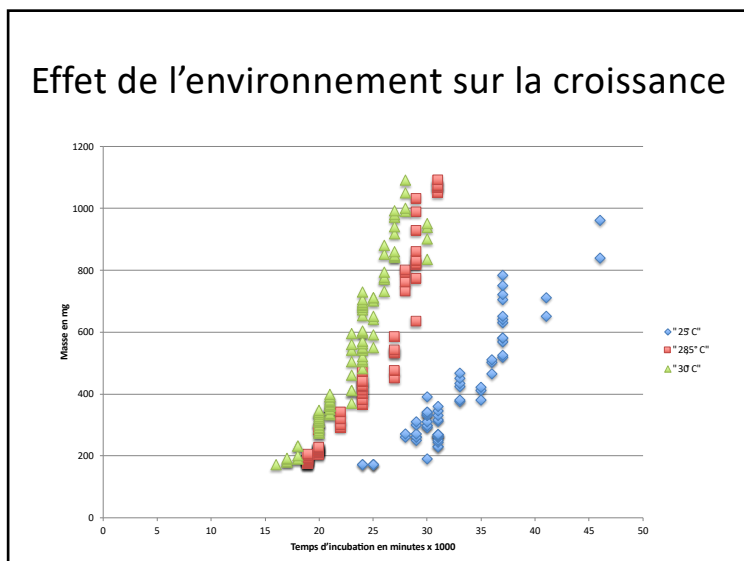
Le modèle de Gompertz suppose que le taux de mortalité est la somme de termes indépendants de l'âge et de termes dépendants de l'âge :

- Gompertz, B., 1825. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. Philosophical Transactions of the Royal Society 115, 513-583.

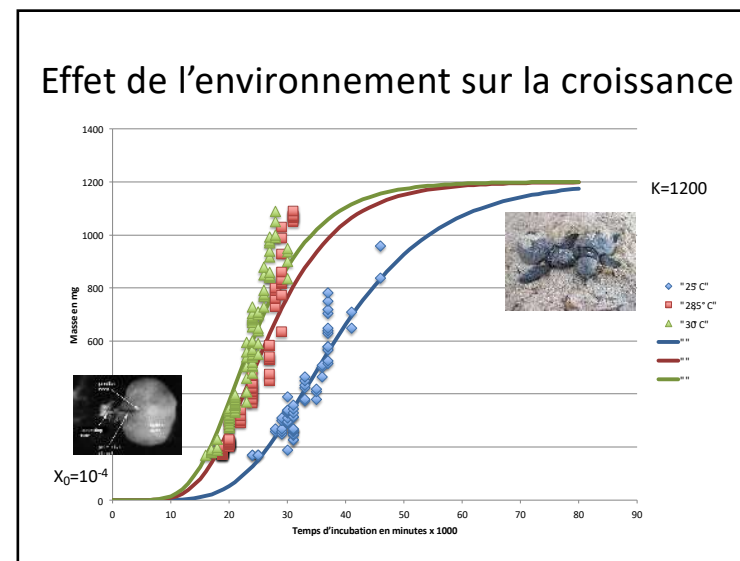
La taille d'un individu ou d'une tumeur est fonction du nombre de cellules :

- Laird, A.K., 1964. Dynamics of tumor growth. British Journal of Cancer 18, 490-502.

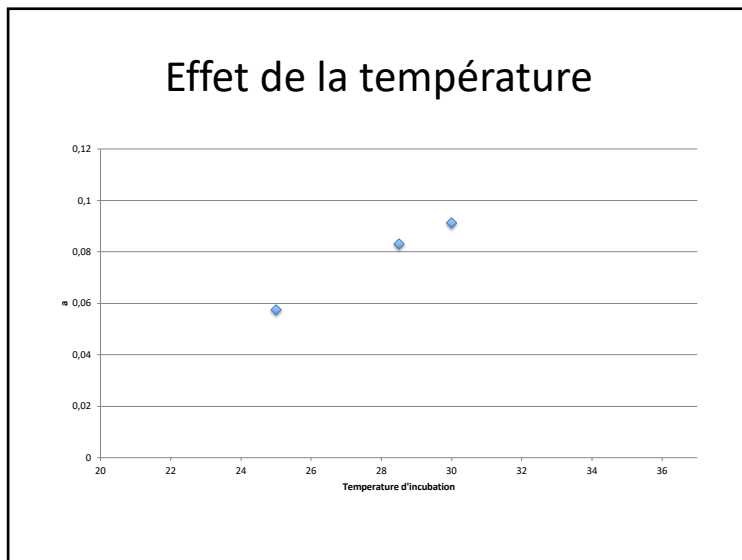
58



59



60



61

Effet de la température sur les réactions biologiques

- **'Reaction rates double for every 10 °C rise in temperature'**
- Le Q_{10} est une mesure de la sensibilité d'un système à un changement de 10 °C ou de 10 K. C'est une mesure sans unité:

$$Q_{10} = \left(\frac{R_2}{R_1}\right)^{10/(T_2-T_1)}$$

- On appelle Q_{10} l'augmentation constatée de vitesse de réaction lorsque l'on augmente la température de 10 °C

Donc le Q_{10} n'est pas indépendant de T !
On oublie donc le Q_{10} !

<http://rod.beavon.org.uk/Q10.htm>

62

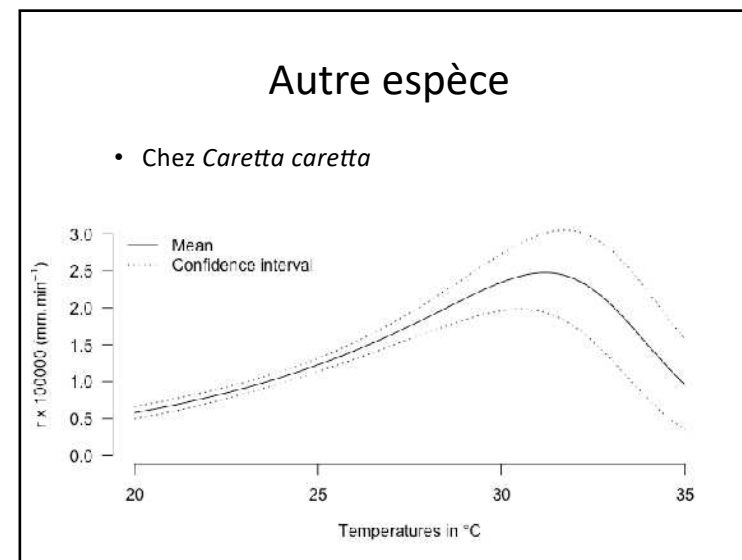
Effet de la température sur les réactions biologiques

- Schoolfield, R.M., P.J. Sharpe, and C.E. Magnuson. 1981. Non-linear regression of biological temperature-dependent rate models based on absolute reaction-rate theory. *Journal of Theoretical Biology* 88 (4):719-31.

$$r(T) = \frac{R_{max} \frac{T}{298} \exp\left[\frac{\Delta H_a}{R} \left(\frac{1}{298} - \frac{1}{T}\right)\right]}{1 + \exp\left[\frac{\Delta H_d}{R} \left(\frac{1}{T} - \frac{1}{T_d}\right)\right]}$$

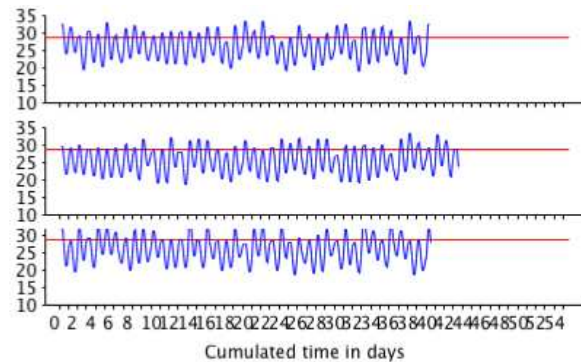
FIG. 1. Arrhenius plot (log rate vs. reciprocal degree Kelvin) of specific growth rate (generations per hour) of *E. coli* in glucose-minimal medium (O'Donovan et al., 1965) (—) = observed mean growth rates at 10 different temperatures. (A) = rates predicted by equation (4); (B) = plot of equation (4) numerator; (C) and (D) = plots of equation (4) numerator divided by second and third term of denominator, respectively. The slopes of lines (B), (C) and (D) are related to enthalpy parameters, as described in the text.

63



64

Une incubation réelle



65

Travailler en pas à pas

- Equation différentielle du modèle dynamique de Gompertz

$$\frac{dx}{dt} = ax \ln\left(\frac{K}{x}\right)$$

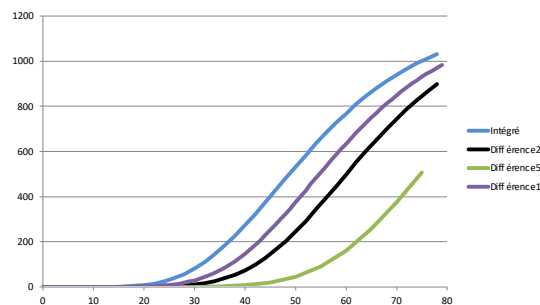
On peut l'écrire sous forme d'une relation de récurrence si $\Delta t \rightarrow 0$

$$\Delta x = ax \ln\left(\frac{K}{x}\right) \Delta t$$

66

Travailler en pas à pas

- Equation différentielle du modèle dynamique de Gompertz appliquée par différence



67

Méthode de Runge-Kutta

- Les méthodes de Runge-Kutta sont des méthodes d'analyse numérique d'approximation de solutions d'équations différentielles.
- Elles ont été nommées ainsi en l'honneur des mathématiciens Carl Runge et Martin Wilhelm Kutta lesquels élaborèrent la méthode en 1895-1901.

Runge, C.D.T., 1895. Über die numerische Auflösung von Differentialgleichungen. *Mathematische Annalen*, Springer 46, 167-178.

Kutta, W., 1901. Beitrag zur näherungsweise Integration totaler Differentialgleichungen. B. G. Teubner, Leipzig, Germany.

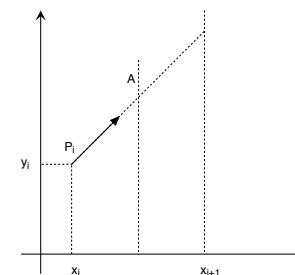
68

Principe de la méthode

- Soit une équation différentielle du premier ordre : $dy = f(x,y) dx$
- La méthode RK4 utilise plusieurs points intermédiaires pour calculer la valeur de y_{i+1} à partir de la valeur de y_i .

69

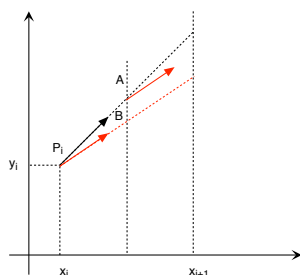
Méthode de Runge-Kutta



Si on s'arrête là, on fait simplement une projection du point en fonction de la dérivée en ce point sans prendre en compte que la courbe n'est pas forcément linéaire entre les points x_i et x_{i+1} . On a vu que c'était très faux si Δt n'est pas très proche de 0.

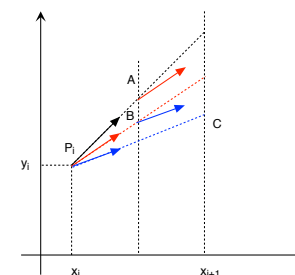
70

Méthode de Runge-Kutta



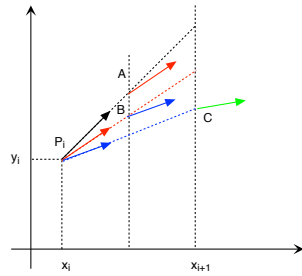
71

Méthode de Runge-Kutta



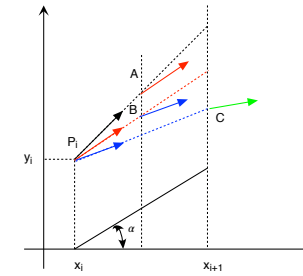
72

Méthode de Runge-Kutta



73

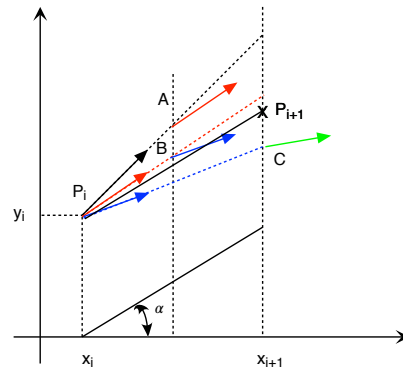
Méthode de Runge-Kutta



$$\text{tg}\alpha = \frac{1}{6} \left[\left(\frac{dy}{dx} \right)_i + 2 \times \left(\frac{dy}{dx} \right)_{(A)} + 2 \times \left(\frac{dy}{dx} \right)_{(B)} + \left(\frac{dy}{dx} \right)_{(C)} \right]$$

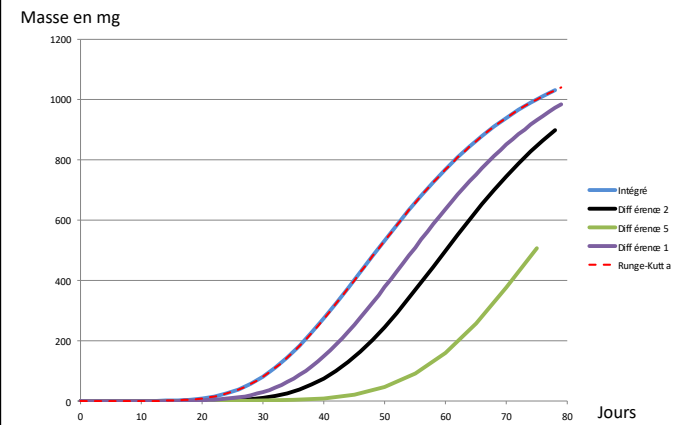
74

Méthode de Runge-Kutta

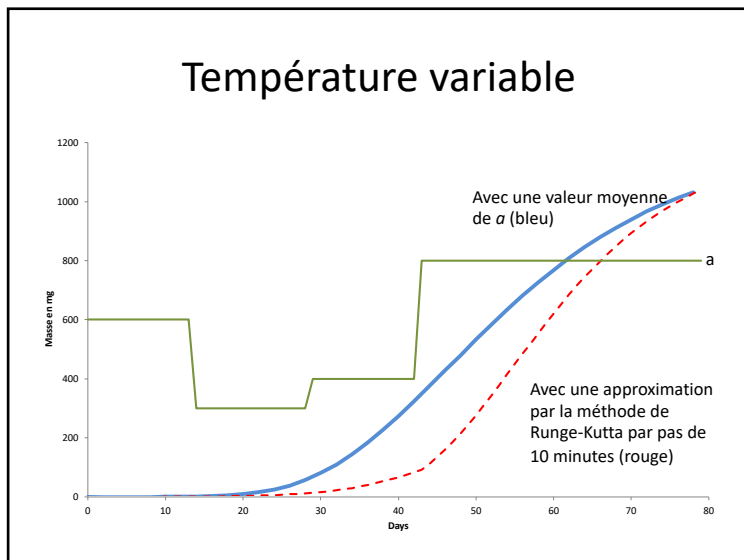


75

Application de la méthode



76



77

Publication

- Girondot M, Kaska Y (2014) A model to predict the thermal reaction norm for the embryo growth rate from field data. *J Therm Biol* 45:96-102

The thumbnail shows the cover of the journal 'Journal of Thermal Biology'. The title of the article is 'A model to predict the thermal reaction norm for the embryo growth rate from field data' by Marc Gironde and Yulia Kaska. The cover includes the journal logo and a small image of a bird.

78