

"Kinetics", mode d'emploi

Généralités

Le programme "Kinetics", une macro complémentaire pour Microsoft Excel (édition 2007 ou ultérieure), vise à rendre accessible à tous la modélisation des cinétiques d'interactions moléculaires et des pharmacocinétiques, ces deux questions étant en fait très proches sur le plan mathématique : résolution de systèmes d'équations différentielles, linéaires ou non-linéaires (bilinéaires ou de type Michaelis-Menten par exemple).

Ce programme, comme ceux de la même série ("Interactions", "Population kinetics") reprend certains concepts d'Equilex (Raguin et al. 2002) : disposition préétablie des données dans une feuille Excel, compilation des formules, utilisation d'Excel pour les affichages graphiques, flexibilité pour tenter de couvrir un maximum de situations expérimentales, utilisation de la feuille Excel comme un tableau de bord permettant de voir rapidement l'ensemble du problème, les conditions de sa résolution et les résultats. En particulier, "Kinetics" permet de regrouper des expériences ou de tenir compte de plusieurs mesures simultanées sans nécessiter de calcul préliminaire de nature à biaiser le résultat (soustraction de bruit de fond, calcul de rapports, etc.).

Rien de très sophistiqué dans les routines de calcul basées sur la boîte à outil "Numerical Recipes in Pascal" de 1989. Pour la résolution des systèmes d'équation différentielles, "Kinetics" se sert des routines classiques de Runge et Kutta (voir Wikipedia pour plus d'information sur ces méthodes) ou de la routine Chu et Berman qui a été réécrite à partir de l'algorithme publié dans l'article de base (Chu et Berman, 1974). Cette routine est spécialement efficace lorsque les solutions des équations différentielles sont des exponentielles (dans ce cas elle est exacte) ou sont proches de fonctions exponentielles (combinaisons linéaires d'exponentielles, par exemple). Dans les problèmes qui nous occupent, c'est la routine à utiliser en priorité. Elle est également utilisée dans Conversational SAAM, SAAM II et WinSAAM (Boston et al. 1981, Barrett et al. 1998, Novotny et al. 2003).

Une interface en Visual Basic (Kinetics.xlam) transfère les informations à une dll (Kinex_32.dll ou Kinex_64.dll) écrite en Pascal qui résout le problème et renvoie la solution vers l'interface qui modifie la feuille de calcul. Visual Basic 32 bits limitant la taille des structures et le volume de transfert de données, celui-ci passe par le disque dur (enregistrement et lecture de fichiers). Ceci ne grève pas la vitesse car la taille des données reste limitée. La dll intègre les routines de calcul décrites ci-dessus et un pseudo-compilateur qui traite toutes les expressions littérales écrites dans la feuille de calcul. Ce pseudo-compilateur est très efficace, même s'il n'est sans doute pas très orthodoxe du point de vue informatique. Le calcul les dérivées de l'erreur d'ajustement par rapport aux paramètres ajustables passe par des répétitions de la simulation des cinétiques. Le programme

implémente l'algorithme de Levenberg-Marquardt (H.P. Gavin, people.duke.edu/~hpgavin/ce281/lm.pdf).

La feuille Excel est organisée en blocs séparés par des cellules vides. Le programme reconnaît les mots clés : **"Initial conditions"**, **"Data"**, etc. écrits dans le coin en haut à gauche de chaque bloc de données. "Kinetics" délimite tout seul les blocs (c'est pour ça qu'il faut laisser des cellules vides entre les blocs) et leur attribue un fond en jolies couleurs pastel.

Initial conditions		F1	Interrupt parameters						
	1	A1		1					
Adjustable parameters		A1	A2	Data	Time	Observed	Observed	F1	
Value		2.03E+01	2.08E-01	1	0				20.339
SD		7.09E-01	1.03E-02	1	1	17.000	16.000		16.522
Equal to				1	2	13.000	14.000		13.421
Adjustable	Y	Y		1	4	9.000	8.500		8.857
				1	5	7.500	7.000		7.194
Equations		F1	Fonction						
d(F1)/dt		-A2			d(F1)/dt = - A2 * F1				
Fit	Error type:	Damping:	Iterations:	Tolerance	Fit tolerance:	Experimental	Lack of fit:		
CHU-BERMAN	2	0.00E+00	1	1.00E-06	1.00E-05	1.04E-01	1.10E-01		

Tous les blocs doivent exister dans la feuille, à l'exception du bloc "Interrupt parameters" qui peut être omis pour cinétiques à une seule étape.

Le programme résout un système d'équations différentielles avec comme variable indépendante le Temps (T). Dans le fichier "Kinetic analyses.xls", la feuille "Monexponentielle" donne un exemple très simple avec utilisation de paramètres ajustables dans la définition des "Sorties" (voir ci-dessous). La feuille "Interaction ligand récepteur" montre comment ça marche pour une interaction moléculaire (non-linéaire) et la feuille "Multiple dosing" pour une pharmacocinétique avec injections multiples. D'autres exemples sont également donnés.

La plupart des entrées inscrites dans les cellules de la feuille de calcul sont compilées par le programme. Presque toutes les cases (sauf les numéros d'interruption, les temps ou les données observées dans le bloc "Data", bien sûr) peuvent contenir une formule écrite en langage quasi naturel (exemple $A1*(F1+F2)$). On peut utiliser les opérations courantes (plus Ln, pour les logarithmes, et Ex pour les exponentielles), y compris parenthèses et puissances : $^$. Toutes ces formules sont compilées en début d'exécution.

Il est pratique d'utiliser un codage simple des différentes variables du système : les F_i sont les solutions du système d'équations différentielles, les P_i sont des paramètres qui peuvent changer à chaque interruption mais ne sont pas ajustables, les A_i sont des paramètres ajustables. Ce codage est utilisé dans ce manuel. Cependant, l'utilisateur peut définir les noms de tous les paramètres et de tous les compartiments à sa guise.

Pour les espèces moléculaires (ou compartiments) F_i , i peut aller de 1 à 30. Pour les paramètres P_i et A_i , i peut aller de 1 à 50. On peut traiter jusqu'à 30 interruptions. Le programme peut calculer pour chaque temps jusqu'à 40 valeurs différentes. Pour chaque temps, on peut entrer jusqu'à 10 répliques (données expérimentales). Ce nombre peut aller de 0 (simulation seulement) à 10 et il n'est pas nécessaire qu'il y ait toujours le même

nombre de répliques. Le nombre de données, toutes interruptions confondues, est limité à 1000.

Le temps est représenté explicitement par T sans indice. Il peut apparaître explicitement dans les formules.

Attention : Excel interprète les formules introduites dans les cases. Par exemple -A1 est compris par Excel comme une référence à la cellule A1. Un message d'erreur sera renvoyé si celle-ci n'est pas numérique. Il faut alors soit forcer le format des cellules à "texte", soit démarrer la formule par une apostrophe ' : '-A1 sera alors bien compris comme du texte que Kinetics saura traiter.

Installation

Kinetics est une add-in qui doit être installée dans Excel : il faut copier les fichiers "kinex_32.dll" et "Kinex_64.dll" dans le répertoire C:\Windows et le fichier "Kinetics.xlam" comme les fichiers d'exemples et autres dans un répertoire au choix. Ensuite il faut lier Excel au complément "Kinetics.xlam" : voir "To install.txt" et aide Excel. Cela fera apparaître un onglet "Expert addins" dans la barre de menu, s'il n'existait pas déjà, et un onglet appelé "Kinetics" dans le ruban.



- Créez un dossier de votre choix et copiez le fichier "Kinetics.xlam" dans celui-ci
- Copiez le fichier "Kinex_32.dll" et "Kinex_64.dll" dans le dossier C:\Windows (nécessite des privilèges d'administrateur)
- Ouvrez Excel, accédez à Options (l'emplacement dépend de la version d'Excel), puis à Compléments et liez le fichier "Kinetics.xlam" comme complément
- Un nouvel onglet "Expert addins" doit apparaître dans le ruban
- En cliquant sur l'onglet, une commande "Kinetics" doit afficher

Les interruptions

Les interruptions reprennent les notions utilisées dans SAAM et WinSAAM. Il s'agit de permettre la simulation d'un système dans lequel certains paramètres sont changés à un instant donné. Exemples : lavage après un certain temps d'incubation ou fin de perfusion. Cela permet aussi de traiter simultanément des expériences différentes qui dépendent de paramètres ajustables communs.

Ces interruptions sont numérotées consécutivement à partir de 1. L'interruption *i* se produit à un instant donné qui précède immédiatement le temps indiqué dans le bloc **"Data"** à la première ligne qui commence par le numéro *i*. À ce moment-là, les conditions initiales, données dans le bloc **"Initial conditions"**, et les valeurs des paramètres P_i , données dans le bloc **"Interrupt parameters"** sont changées. Noter que ces valeurs peuvent être littérales.

Initial conditions	F1	F2	F3		Interrupt par P1
1	2.00E+01	1.00E+01	0.00E+00		1 1.00E+00
2	0.00E+00	F2	F3		2 1.00E+00

Le numéro d'interruption, 1, 2..., fait le lien entre les blocs **"Initial conditions"**, **"Interrupt parameters"** et **"Data"**. C'est dans ce dernier bloc qu'est donné le temps auquel se produit l'interruption.

Dans le bloc **"Initial conditions"**, on donne les valeurs des variables F1, F2, F3... aux différentes interruptions (la première correspond au temps 0 de l'expérience). Ces valeurs initiales peuvent être numériques ou être des fonctions des paramètres fixes et des variables F_i . Ainsi, si une variable F_i est inchangée à l'interruption 2, on met F_i dans la case pour indiquer que la valeur de F_i après l'interruption est la même qu'avant. Dans l'exemple "Interaction ligand récepteur", c'est le cas pour F2 et F3, tandis que F1 est mis à 0 au temps 5 minutes. Ce système permet de modéliser les lavages, fins de perfusion, injections d'un autre produit ou d'une autre dose, ajouts d'un compétiteur...

Si on a besoin de paramètres pour décrire les changements lors des interruptions, alors on se sert du bloc **"Interrupt parameters"**. Ce bloc a obligatoirement le même nombre de lignes que le bloc **"Initial conditions"**, puisque chaque ligne reflète une interruption (encore une fois l'interruption 1 est le début de la cinétique, elle peut bien sûr être la seule).

Data

Le bloc Data regroupe les entrées-sorties, c'est-à-dire les données observées et les résultats calculés pour chaque temps

Data	Time	Observed	Observed	F1/A1	F1	F2/A5	F2	F3	F4	F1+F2
1	0.0			27.05%	1.000					1
1	0.7	27.0%	26.5%	26.14%	0.966	1.03%	0.0	0.7	0.0	0.98
1	1.6	24.5%	24.6%	25.04%	0.926	2.24%	0.0	1.6	0.0	0.96
1	2.0			24.62%	0.910	2.69%	0.0	1.9	0.0	0.95
1	2.5			24.06%	0.890	3.28%	0.1	2.4	0.1	0.94
1	2.9	23.5%	23.5%	23.65%	0.874	3.71%	0.1	2.7	0.1	0.93
1	3.0			23.53%	0.870	3.84%	0.1	2.8	0.1	0.93
1	6.0			20.70%	0.765	6.61%	0.1	5.2	0.4	0.87
1	12.0			16.55%	0.612	9.93%	0.2	9.3	1.2	0.77
1	18.0			13.72%	0.507	11.35%	0.2	12.7	2.2	0.69
1	19.2			13.26%	0.490	11.49%	0.2	13.3	2.4	0.67
1	21.0	12.7%	12.6%	12.62%	0.467	11.63%	0.2	14.2	2.8	0.65
1	24.0			11.71%	0.433	11.70%	0.2	15.5	3.3	0.62
1	48.0			7.27%	0.269	9.52%	0.2	23.6	7.5	0.42
1	92.9	3.5%	3.6%	3.58%	0.132	4.94%	0.1	32.2	12.6	0.21
1	124.2	2.2%	2.2%	2.22%	0.082	3.06%	0.0	35.5	14.6	0.13
1	360.0			0.06%	0.002	0.08%	0.0	40.7	17.7	0.00
1	720.0			0.00%	0.000	0.00%	0.0	40.9	17.8	0.00

La première colonne du bloc correspond au numéro de l'interruption. La deuxième colonne du bloc Data est le temps. On peut trouver, dans cette

colonne des temps identiques s'ils n'appartiennent pas à la même interruption, mais pour une interruption donnée, ils doivent être strictement croissants.

Viennent ensuite les "Sorties" qui représentent ce qu'on mesure dans l'expérience ou ce qu'on veut simuler. Les données mesurées expérimentalement sont listées dans le bloc Data dans les colonnes "Observed" Il peut y avoir plusieurs colonnes "Observed". Chaque "Sortie" est définie dans une colonne, à droite des données expérimentales, qui ne commence pas par "Observed", mais par l'expression analytique de la mesure en fonction des F1, F2... quantités de produit dans le compartiment 1, 2, etc. et paramètres pour chaque temps. À noter qu'il n'est pas nécessaire d'avoir des valeurs expérimentales ("Observed").

Équations

Les équations du système sont données dans le bloc "Equations".

Equations	F1	F2	Fonction
d(F1)/dt	-A1-A3	A2	P1
d(F2)/dt	A1	-A2	

Il faut comprendre que les colonnes et les lignes (en nombre égal en dessous de la ligne de titres) représentent une matrice M. La dernière colonne à droite ("Function" ou "Fonction") est un vecteur Y.

Le vecteur des dérivées par rapport au temps dF/dt est calculé comme $M \cdot F + Y$ en notation matricielle. Autrement dit, pour l'espèce F_i , $dF_i/dt = \sum M_{ij} F_j + Y_i$. En analyse compartimentale, les composants F_i sont échangés entre les compartiments j et i proportionnellement à la valeur dans la source. La méthode de résolution "Chu-Berman" est basée sur cette notation matricielle. La fonction Y_i reçoit toute contribution à dF_i/dt qui n'est proportionnelle à aucun des F_j . Une constante peut, par exemple, représenter une perfusion dans un compartiment. Le temps peut apparaître explicitement dans ces équations. Pour traiter des échanges non-linéaires, il suffit d'introduire la non-linéarité dans les valeurs de la matrice M (voir exemples).

Paramètres

Les paramètres sont de deux types :

1. les paramètres ajustables, A_j , que le programme peut estimer s'il dispose de données expérimentales en nombre suffisant.

Adjustable parameters	A1	A2	A3
Value	5.00E-01	3.00E-01	2.00E-01
SD	NA	NA	NA
Equal to			
Adjustable	N	N	N

2. les autres paramètres, P_i , qui peuvent changer au cours de l'expérience à certains temps donnés que l'on appelle des interruptions.

Interrupt parameters	P1	
	1	5.00
	2	0.00
	3	5.00
	4	0.00
	5	5.00
	6	0.00

Les définitions de ces deux types de paramètres sont données dans les blocs "**Adjustable parameters**" et "**Interrupt parameters**". Pour les ajustables, il faut donner une valeur initiale, et indiquer s'il faut les ajuster ou pas (Y or N). Le programme renverra une estimation avec une estimation asymptotique du SD s'ils sont ajustables. On peut forcer ces paramètres à une valeur quelconque (fonction écrite dans la ligne "Equal to"). Pour les Pi, la première colonne donne le numéro de l'interruption, les autres le nom, Pi, du paramètre et sa valeur à chaque interruption (valeur numérique ou littérale, les définissant en fonction des autres données du problème, Aj par exemple).

Le bloc Fit

Le dernier bloc est le bloc "**Fit**". Il n'y a pas en général à y toucher, sauf pour le nombre d'itérations : si on met 0, le programme fait seulement une simulation, si on met 3, 5 ou 10, il y aura autant d'itérations pour identifier les paramètres ajustables. Le fit est obtenu lorsqu'une valeur de SD est renvoyée pour chaque paramètre ajustable.

Fit	Error type:	Damping:	Iterations:	Tolerance	Fit tolerance:	Experimental	Lack of fit:
CHU-BERMAN	1	0.00E+00	0	1.00E-04	1.00E-03	0.00E+00	0.00E+00

On retrouve l'amortissement ("Damping") de Marquardt (indispensable quand il y a de fortes corrélations entre paramètres ajustables) et les tolérances sur les calculs de Fi et sur l'ajustement final, etc.

Un autre choix intéressant est celui de la méthode de résolution du système d'équations différentielles. CHU-BERMAN est une méthode spécialement développée pour la pharmacocinétique par Chu et Berman au NCI à partir des années 60. C'est la méthode phare de SAAM et CONSAAM. On peut aussi utiliser les classiques méthodes de Runge et Kutta au 4^{ème} ou au 5^{ème} ordre : RUNGE-KUTTA 4 ou 5. En principe toutes doivent donner le même résultat avec des différences de vitesse de calcul suivant les problèmes. Elles dépendent toutes du paramètre "Tolerance" qui ne doit être ni trop grand, ni trop petit (entre 10^{-3} ou 10^{-6}).

Exécution

Lorsque la feuille est complète, on lance le programme à partir du ruban Excel. Une petite fenêtre apparaît donnant le choix entre exécution (Calculate) ou affichage d'une aide (Aide ou Help). Cette fenêtre disparaît pendant que le programme tourne et la feuille Excel est réactualisée à chaque itération. Le nombre d'itérations effectuées et le temps écoulé sont affichés dans la barre "Status" (en bas à gauche de la feuille Excel). On ne

peut rien faire dans Excel tant que la petite fenêtre n'est pas réapparue. À ce moment, on peut relancer le calcul, consulter l'aide ou fermer la fenêtre.

Conclusion

L'idée était de développer une plate-forme de modélisation cinétique la plus flexible possible grâce à l'usage de formules dans les cases de la feuille Excel et non pas une boîte noire traitant certains systèmes prédéfinis. Les exemples fournis donnent une idée des possibilités.

Il est possible que des erreurs subsistent. Prière de les signaler, comme toute autre question relative à Kinetics à : barbet@arronax-nantes.fr.

Ce logiciel est gratuit pour l'utilisation et la distribution à condition que ceux-ci ne soient pas à but lucratif, que le GIP ARRONAX soit remercié dans la communication ou la publication de tout résultat obtenu en l'utilisant et que les bénéficiaires du transfert s'engagent aux conditions précédentes.

Les sources peuvent être obtenues sur demande expresse à barbet@arronax-nantes.fr.

Références

Raguin O, Gruaz-Guyon A, Barbet J. Equilibrium expert: an add-in to Microsoft Excel for multiple binding equilibrium simulations and parameter estimations. *Anal Biochem.* 2002; 310: 1-14.

Numerical Recipes in Pascal (First Edition): The Art of Scientific Computing, Cambridge University Press, October 27, 1989.

Chu SC, Berman M. An exponential method for the solution of systems of ordinary differential equations. *Comm. ACM.* 1974; 17: 699-702.

Boston RC, Greif PC, Berman M. Conversational SAAM--an interactive program for kinetic analysis of biological systems. *Comput Programs Biomed.* 1981; 13:111-9.

Barrett PH, Bell BM, Cobelli C, Golde H, Schumitzky A, Vicini P, Foster DM. SAAM II: Simulation, Analysis, and Modeling Software for tracer and pharmacokinetic studies. *Metabolism.* 1998; 47: 484-92.

Novotny JA, Greif P, Boston RC. WinSAAM: application and explanation of use. *Adv Exp Med Biol.* 2003; 537:343-51.