

# MODÉLISATION STOCHASTIQUE ET ESTIMATION DE LA CROISSANCE TUMORALE

Modibo DIABATE <sup>1</sup> & Loren COQUILLE <sup>2</sup> & Adeline SAMSON <sup>3</sup>

<sup>1</sup>*Laboratoire Jean Kuntzmann (LJK), Université Grenoble Alpes  
(modibo.diabate@univ-grenoble-alpes.fr)*

<sup>2</sup>*Institut Fourier (IF), Université Grenoble Alpes  
(loren.coquille@ujf-grenoble.fr)*

<sup>3</sup>*Laboratoire Jean Kuntzmann (LJK), Université Grenoble Alpes  
(adeline.leclercq-samson@imag.fr)*

**Résumé.** L'objectif principal de ce projet est d'aider à la compréhension de la résistance des tumeurs aux approches immunothérapeutiques. Elle consiste à étudier un modèle stochastique rendant compte des phénomènes qui se déroulent pendant l'immunothérapie du cancer de la peau et à estimer (avec des données biologiques) les paramètres du modèle en utilisant des modèles non linéaires à effets mixtes (NLMEM) et l'algorithme SAEM (Stochastic Approximation Expectation Maximization). Nous disposons d'une base de données de mesures de croissance tumorales recueillies chez 19 souris, réparties en trois groupes (sans traitement, avec immunothérapie, avec immunothérapie et restimulation). Nous avons, dans un premier temps, estimé (et validé) les paramètres associés à la limite déterministe du modèle stochastique. En revenant au modèle stochastique avec ces paramètres, nous estimons la probabilité d'extinction des cellules T. Cette extinction (qui peut être interprétée comme une rechute de la maladie) est un événement rare dont l'estimation de la probabilité de survenue implique l'utilisation de techniques spécifiques comme l'algorithme d'importance Splitting.

**Mots-clés.** Modélisation stochastique, Estimation de paramètres, Modèles non linéaires à effets mixtes, algorithme SAEM, Données censurées, Cancer, Résistance des tumeurs, Immunothérapie, Estimation de probabilité d'évènements rares, Importance Splitting, . . .

**Abstract.** The main objective of this project is to help in the understanding of cancer resistance to immunotherapeutic approaches. It consists in the study of a stochastic model for various phenomena arising in immunotherapy of skin cancer; and the estimation (with biological data) of the model parameters using Non Linear Mixed Effects Models (NLMEM) and the Stochastic Approximation Expectation Maximization (SAEM) algorithm. We have a database of tumor growth measurements for 19 mice divided into three groups (without treatment, with immunotherapy, with immunotherapy and restimulation). We first, estimated (and validated) the parameters associated to the (deterministic) large population approximation of the model. Heading back to the stochastic model with these parameters, we estimate the probability of extinction of T cells. This extinction (which can be interpreted as a relapse of the disease) is a rare event whose probability estimation involves the use of specific techniques such as the algorithm of importance Splitting.

**Keywords.** Stochastic modeling, Parameter estimation, Non Linear Mixed Effects Models (NLMEMs), SAEM algorithm, Censored Data, Cancer, Tumor resistance, Immunotherapy, Rare event probability estimation, Importance Splitting, ...

## 1 Introduction

Nous considérons dans ce projet l'exemple des mélanomes (tumeurs associées au cancer de la peau) sous une thérapie avec des cellules T. Notre projet s'inscrit dans la continuité des travaux de Loren Coquille et ses collègues (de l'université de Bonn) sur leur modèle stochastique pour l'immunothérapie du cancer (M. Baar, L. Coquille et al (2015)) basé sur les expériences de J. Landsberg et al (2012), qui étudient les mélanomes chez la souris sous le traitement appelé Adoptive Cell Transfer (ACT) therapy. Cette approche thérapeutique implique la stimulation de cellules T qui reconnaissent un type spécifique de cellules tumorales de mélanomes (cellules différenciées) par l'intermédiaire de marqueurs spéciaux sur leur surface et qui sont capables de tuer ces cellules de mélanomes différenciées. La thérapie induit une inflammation et les cellules de mélanomes réagissent à cette modification de l'environnement en changeant leur type, c'est-à-dire en passant d'une cellule de mélanomes différenciée à une cellule dédifférenciée (les marqueurs spéciaux disparaissent sur la surface cellulaire). Ce changement de type cellulaire est renforcé si des cytokines pro-inflammatoires, appelés  $TNF_\alpha$  sont présents. On note que le changement de type est réversible. On note également que l'état de la tumeur est un mélange de cellules différenciées et dédifférenciées.

Pendant le traitement du cancer, la rechute est l'un des principaux problèmes auxquels il faut faire face. Dans ce projet, nous nous intéressons à la rechute due à l'épuisement des cellules T qui ne parviennent plus à tuer efficacement les cellules de mélanomes (on note qu'ici, les cellules T peuvent tuer uniquement les cellules différenciées). Ce problème est résolu par une nouvelle stimulation des cellules T qui conduit à un retard dans la survenue de la rechute. Le modèle stochastique sous la forme d'un modèle individu-centré (M. Baar, L. Coquille et al (2015)), dans la limite des grandes populations, converge vers un système dynamique quadratique contenant quelques paramètres (taux de reproduction, de mort et de changement de type).

Considérant le travail de M. Baar, L. Coquille et al (2015) qui soutient qu'il existe des paramètres biologiquement raisonnables pour lesquels la rechute décrite apparaît avec une probabilité raisonnable, notre objectif est d'obtenir une estimation statistique de ces paramètres avec des données biologiques fournies par les auteurs de J. Landsberg et al (2012).

### Données expérimentales

Les données utilisées proviennent de l'hôpital universitaire de Bonn. Trois groupes de souris sont considérés dans cette expérience de l'ACT therapy : le groupe de souris non

traitées qui servent d'individus témoins (groupe CTRL); le groupe de souris qui reçoivent une seule dose de stimulant de cellules T à un temps  $t_a$  au cours de la thérapie (groupe ACT); et le groupe de souris qui reçoivent deux doses de stimulant de cellules T aux temps  $t_a$  (stimulation) et  $t_b$  (re-stimulation) pendant la thérapie (groupe ACT+Re) (J. Landsberg et al (2012)).

La base de données utilisée contient des données pour 19 souris : 5 souris du groupe CTRL (avec un nombre moyen d'observations par souris égal à 10); 7 souris du groupe ACT (avec un nombre moyen d'observations égal à 26); et 7 souris du groupe ACT+Re (avec un nombre moyen d'observations égal à 33). Les informations disponibles sur ces souris sont : le numéro d'identification, la taille du diamètre de la tumeur, la date de mesure de la taille de la tumeur, l'identification du groupe de la souris.

**Remarque :** Les mesures de la taille de la tumeur ne sont pas très précises pendant les premières semaines de l'évolution de la tumeur (en raison des difficultés dans la détection des tumeurs à ce stade) et pour une taille de tumeur supérieure ou égale à 10 mm la souris est tuée. Pour ces raisons, nous travaillons avec des données censurées.

## 2 Modèle pour l'étude de la croissance tumorale

Le modèle utilisé pour l'étude de la croissance des tumeurs est un processus stochastique de naissance et de mort combiné avec un système prédateur-proie (entre les cellules T et les cellules de mélanome différenciées) qui tient compte des changements de type entre cellules différenciées et dédifférenciées (voir Baar, Coquille et al (2015)). Pour l'estimation des paramètres, nous nous intéressons à la limite des grandes populations de ce modèle stochastique. Cette limite correspond à un système dynamique quadratique contenant quelques paramètres (taux de reproduction, de mort et de changement de type).

**On note :**

- $M(t)$  : la population de cellules différenciées à l'instant  $t$
- $D(t)$  : la population de cellules dédifférenciées à l'instant  $t$
- $T(t)$  : la population de cellules T à l'instant  $t$
- $A(t)$  : la population de  $TNF_\alpha$  à l'instant  $t$

L'évolution des populations de cellules impliquées dans ce traitement ACT-therapy est modélisée par le processus stochastique à quatre dimensions  $Z(t) = (M(t), D(t), T(t), A(t))$ . La limite des grandes populations de ce modèle nous conduit au système ci-dessous :

$$\begin{cases} \dot{\mathbf{n}}_M = (b_M - d_M)\mathbf{n}_M - t_T \mathbf{n}_T \mathbf{n}_M - s_{MD} \mathbf{n}_M + s_{DM} \mathbf{n}_D - s_A \mathbf{n}_A \mathbf{n}_M \\ \dot{\mathbf{n}}_D = (b_D - d_D)\mathbf{n}_D + s_{MD} \mathbf{n}_M - s_{DM} \mathbf{n}_D + s_A \mathbf{n}_A \mathbf{n}_M \\ \dot{\mathbf{n}}_T = -d_T \mathbf{n}_T + b_T \mathbf{n}_M \mathbf{n}_T \\ \dot{\mathbf{n}}_A = -d_A \mathbf{n}_A + l_A^{prod} b_T \mathbf{n}_M \mathbf{n}_T \end{cases}$$

où  $n_X$  (avec  $X = M, D, T$  ou  $A$ ) représente le nombre de cellules  $X$  à l'instant  $t$  dans le modèle déterministe et  $\dot{n}_X$  représente la variation de ce nombre en fonction du temps.

Les paramètres  $b_M, b_D, b_T, l_A^{prod}, d_M, d_D, t_T, d_T, d_A, s_A, s_{MD}, s_{DM}$  sont à estimer.  $b_M, b_D$  cor-

respondent respectivement aux taux de naissance des cellules différenciées et des cellules dédifférenciées. Ces taux représentent (dans le modèle stochastique) le temps d’attente avant d’observer la naissance d’une cellule différenciée (respectivement, une cellule dédifférenciée).  $d_M, d_D$  correspondent respectivement aux taux de mort des cellules différenciées et des cellules dédifférenciées ;  $s_{MD}, s_{DM}$  correspondent aux taux de changement de type entre cellules différenciées et dédifférenciées ;  $b_T$  et  $d_T$  correspondent respectivement aux taux de naissance et de mort (naturelles) des cellules T ;  $t_T$  correspond à un taux de mort des cellules différenciées sous l’effet de l’immunothérapie ;  $d_A$  correspond au taux de mort des messagers chimiques  $TNF_\alpha$  ;  $l_A^{prod}$  correspond à un taux de production des  $TNF_\alpha$  induite par l’immunothérapie ; enfin,  $s_A$  correspond à un taux de changement de type des cellules différenciées vers les cellules dédifférenciées induit par l’immunothérapie.

### 3 Estimation de paramètres

Pour l’estimation des paramètres du modèle (déterministe), nous utilisons un modèle non linéaire à effets mixtes. L’estimation de paramètres avec un tel modèle est critique puisque la fonction de vraisemblance n’a généralement pas une forme explicite. Dans ce contexte, l’algorithme SAEM (Stochastic Approximation Expectation Maximization) apparaît être un très bon outil d’estimation.

Nous utilisons une version de l’algorithme SAEM (voir Delyon et al (1999); Monolix (2014)) appelée SAEM-MCMC qui combine l’algorithme SAEM avec une procédure MCMC en tenant compte de la présence des données censurées dans la fonction de vraisemblance du modèle (Samson et al (2006)).

Nous obtenons une estimation des paramètres individuels et de population pour chacun des trois groupes de souris (les paramètres liés au traitement sont nuls chez les souris du groupe CTRL qui n’ont que la phase de croissance de la tumeur). La figure 1 représente l’évolution du diamètre de la tumeur en fonction du temps (obtenue avec les paramètres estimés) pour 4 souris du groupe ACT+Re.

Afin de valider ces résultats, nous calculons les erreurs standards associées aux paramètres à travers une estimation de la matrice de Fisher.

## 4 Étude des rechutes

### 4.1 Extinction des cellules T

Le phénomène qui nous intéresse particulièrement dans cette étude est l’épuisement des cellules T (qui ne parviennent plus à tuer les cellules différenciées de mélanomes durant la thérapie) entraînant une rechute de la maladie (croissance exponentielle de la population des cellules différenciées). Nous assimilons cet épuisement à une extinction de la population des cellules T dans notre modèle. Afin de simuler cette extinction des cellules T, nous utilisons dans le modèle stochastique de naissance et de mort les paramètres estimés à partir du modèle déterministe en nous assurant que les propriétés théoriques de ces paramètres

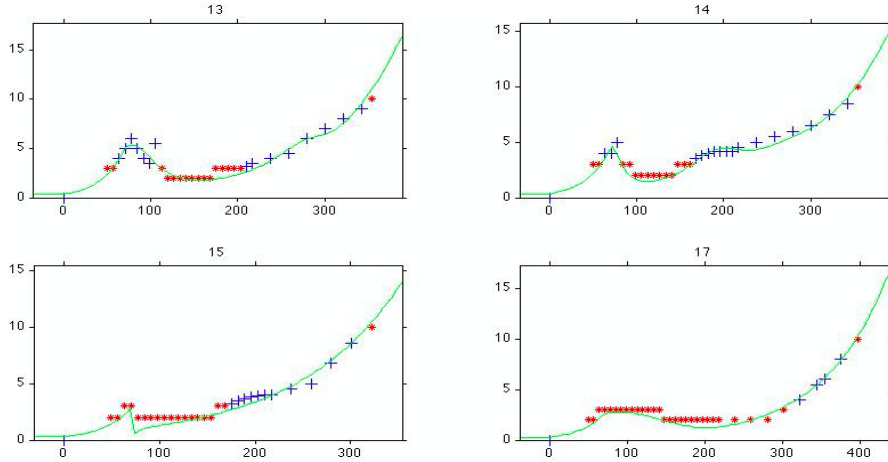


FIGURE 1 – Représentation graphique de l’évolution de la taille de la tumeur (mm) en fonction du temps (jour) pour 4 souris du groupe ACT+Re. Les points bleus représentent les observations non censurées, les points rouges représentent les observations censurées, la ligne verte représente l’évolution de la taille de la tumeur.

permettent d’observer une extinction (avec une probabilité raisonnable). Ainsi, à partir des paramètres estimés dans le modèle déterministe, nous simulons l’évolution des différentes populations impliquées dans l’étude à l’aide du modèle stochastique correspondant. Ces simulations ont permis d’observer que pour certains paramètres, le phénomène d’extinction des cellules T était un évènement peu fréquent ( $< 1\%$ ). Nous nous intéressons à l’estimation de la probabilité d’extinction des cellules T.

## 4.2 Calcul de la probabilité d’extinction des cellules T

Nous procédons par importance splitting en utilisant l’algorithme d’estimation de probabilité d’évènements rares I2S (Importance Splitting and Sampling) qui est une combinaison des méthodes de Splitting et de Sampling adaptée aux processus de Markov décrit dans Jacquemart-Tomi et al (2014). Cette méthode est une version stochastique améliorée de la méthode Importance Splitting de Morio et al (2010). I2S donne une meilleure précision pour l’estimation de la probabilité (meilleures erreurs relatives) par rapport à des méthodes comme la méthode de Monte Carlo ou l’importance Sampling.

Avec cet algorithme I2S, nous calculons progressivement la probabilité que la population des cellules T (représentée par le processus stochastique  $T_t$ ) s’éteigne. Pour cela, on lance plusieurs simulations et on choisit à chaque itération de l’algorithme les trajectoires les plus ”rapides”, c’est à dire, celles qui ont plus de chance d’atteindre le seuil  $S$  (valeur correspondant à l’extinction des cellules T) avant le temps  $t_S$  (premier temps que la trajectoire touche le seuil  $S$ ,  $t_S = \inf\{t \geq 0, T_t \leq S\}$ ). L’évènement  $\{t_S \leq t_F\} = \{T_t \leq S, \text{ pour } t \leq t_F\}$  avec  $t_F$  un temps d’arrêt aléatoire, est l’évènement rare. Sa probabilité  $\mathbb{P}(t_S \leq t_F)$  est définie

(selon le principe du splitting) par

$$\mathbb{P}(t_S \leq t_F) = \prod_{k=1}^m p_k \text{ avec } p_1 = \mathbb{P}(t_1 \leq t_F) \text{ et } p_k = \mathbb{P}(t_k \leq t_F \mid t_{k-1} \leq t_F) \text{ pour } k = 2, \dots, m$$

## 5 Conclusion

Dans cet exposé, nous introduisons un modèle stochastique de résistance à l'immunothérapie. Nous montrerons les résultats d'estimation des paramètres à partir de données longitudinales en utilisant le modèle déterministe, limite en grande population du modèle stochastique. Enfin, nous montrerons comment, en utilisant des algorithmes d'estimation de probabilité d'événements rares pour des processus en temps continu, nous pouvons estimer la probabilité de rechute après immunothérapie.

## Bibliographie

- [1] M. Baar, L. Coquille, H. Mayer, M. Hölzel, M. Rogava, T. Tüting, and A. Bovier. A stochastic individual-based model for immunotherapy of cancer. *Scientific Reports*, 2015.
- [2] J. Landsberg, J. Kohlmeyer, M. Renn, T. Bald, M. Rogava, M. Cron, M. Fatho, V. Lennerz, T. Wölfel, M. Hölzel, et al. Melanomas resist T-cell therapy through inflammation-induced reversible dedifferentiation. *Nature*, 490(7420) :412–416, 2012.
- [3] D. Jacquemart-Tomi, J. Morio, F. Le Gland. A combined importance splitting and sampling algorithm for rare event estimation. Raghu Pasupathy ; Seong-He Kim ; Andreas Tolk ; Ray R. Hill ; Mike E. Kuhl. *Proceedings of the 2013 Winter Simulation Conference, Washington 2013, Dec 2013, WASHINGTON, United States*.
- [4] B. Delyon, M. Lavielle, and E. Moulines. Convergence of a stochastic approximation version of the EM algorithm. *Annals of statistics*, pages 94–128, 1999.
- [5] S. N. Ethier and T. G. Kurtz. *Markov processes : characterization and convergence*, volume 282. John Wiley & Sons, 2009.
- [6] A. Samson, M. Lavielle, and F. Mentré. Extension of the SAEM algorithm to left-censored data in nonlinear mixed-effects model : Application to HIV dynamics model. *Computational Statistics and Data Analysis*, 51(3) :1562–1574, 2006.
- [7] J. Morio, R. Pastel and F. Le Gland. An overview of importance splitting for rare event simulation. *European Journal of Physics*, 2010
- [10] S. Barthelmé and N. Chopin. Expectation propagation for likelihood-free inference. *Journal of the American Statistical Association*, 109(505) :315–333, 2014.
- [11] Monolix. *Monolix Users Guide Version 4.3.3, A software for the analysis of nonlinear mixed effects models*, 2014.