La détection de non-linéarité par titrage du bruit : encore une technique dépendant du choix de l'observable.

Elise Roulin, Ubiratan Santos Freitas & Christophe Letellier

CORIA UMR 6614 — Université et INSA de Rouen, BP. 12, 76801 Saint-Etienne du Rouvray cedex roulin@coria.fr

Résumé. L'identification de la chaoticité d'une dynamique expérimentale demeure un véritable défi, principalement parce qu'elle requiert la preuve formelle d'un déterminisme sous-jacent. Même si des modèles déterministes ont déjà été obtenus à partir de données expérimentales, il est rare de pouvoir fournir cette preuve de façon satisfaisante dans le cas de systèmes biologiques. Aussi, une alternative consiste à détecter la présence ou non d'un processus non-linéaire gouvernant la dynamique. La technique de « titrage du bruit » développée par Barahona et Poon repose sur la comparaison entre les prédictions d'un modèle linéaire et celles issues d'un modèle non-linéaire. Elle est relativement efficace, à condition qu'elle soit utilisée dans de bonnes conditions, c'est-à-dire à partir de données correctement échantillonnées et avec des modèles à la structure adaptée. À l'aide du système de Rössler, nous montrons qu'il existe une dépendance des résultats au choix de la variable utilisée.

Abstract. Identifying chaotic dynamics from biological data is very challenging, mainly because it requires a conclusive proof for an underlying determinism. Even if deterministic models were already found from experimental data sets, it is rarely possible to provide such a proof for biological systems. An alternative consists in detecting a nonlinear process governing the dynamics under investigations. The noise titration technique, based on the comparison between one-step-ahead predictions by linear and nonlinear models, is appropriate for such a task. But it has to be used in right conditions, that is, to be applied on well sampled data and using models with appropriate structures. The noise titration technique is applied to the Rössler system to show that the detection of nonlinear component is also related to the concept of observability.

1 Introduction

De nombreuses tentatives ont été menées afin de développer des techniques capables d'identifier de manière indiscutable l'existence d'un comportement chaotique et ce, uniquement à partir d'une seule série temporelle. Dans le domaine de la biologie, les enjeux liés à cette question sont nombreux, aussi bien d'un point de vue physiologique que thérapeutique. Toutefois, seule la sensibilité aux conditions initiales est le plus souvent testée (plus grand exposant de Lyapunov, etc.) et la question du déterminisme est rarement abordée. À ce jour, la seule preuve satisfaisante de l'existence d'un déterminisme sous-jacent serait l'obtention d'un modèle global reproduisant la dynamique et vérifiant par ailleurs la propriété de sensibilité aux conditions initiales. Mais un tel modèle n'a encore jamais été obtenu à partir d'enregistrements biologiques [1,2].

Pour pallier ce manque, Poon & Barahona ont proposé une technique de titrage de bruit de manière à extraire la chaoticité de la dynamique sous-jacente à une série temporelle [3]. Malheureusement, cette technique s'est révélée incapable de distinguer, dans certains cas, une dynamique aléatoire d'une dynamique déterministe [2]. Un objectif plus modeste consiste donc à analyser l'impact de la nonlinéarité sur les propriétés du système étudié, comme nous l'avons déjà fait à partir d'enregistrements électrocardiographiques de patients souffrant de différentes pathologies cardiaques [1]. Cette technique est par exemple utilisée pour distinguer des dynamiques respiratoires spontanées [4] ou assistées mécaniquement [5]. De manière générale, les paramètres définissant la structure des modèles sont laissés libres, ce qui peut biaiser certains résultats. Nous montrons ici que cette technique doit être appliquée avec certaines précautions.

158 E. Roulin, U. S. Freitas & C. Letellier

2 Principe de la détection de non-linéarité

Soit une sérite temporelle $\{s_k\}_{k=1}^n$ correspondant aux données mesurées. Le test de non-linéarité appliqué à la série temporelle $\{s_k\}$ consiste à comparer l'erreur de prédiction à un pas en avant du meilleur modèle linéaire et du meilleur modèle non-linéaire, obtenus tous deux à partir des valeurs de s_k [6]. Les modèles utilisés sont des modèles polynomiaux auto-regressifs, qu'ils soient linéaires ou non. Par exemple, un modèle linéaire d'ordre 5 (décalage temporel maximum autorisé), s'écrit :

$$S_k = \theta_0 + \theta_1 s_{k-1} + \theta_2 s_{k-2} + \theta_3 s_{k-3} + \theta_4 s_{k-4} + \theta_5 s_{k-5}.$$
 (1)

où l'ensemble des valeurs θ_i désigne les paramètres du modèle. Un modèle non-linéaire auquel ce meilleur modèle linéaire peut être comparé, et qui comporterait également six termes, peut s'écrire :

$$S_k = \theta_0 + \theta_1 s_{k-1} + \theta_2 s_{k-2} + \theta_3 s_{k-1} s_{k-2} + \theta_4 s_{k-1}^2 + \theta_5 s_{k-2}^2.$$
(2)

où l'ensemble des valeurs θ_i représente les paramètres du modèle non-linéaire. Le modèle linéaire (1) est d'ordre 5 et de degré d = 1 : il comporte N = 6 termes. Le modèle non linéaire (2) est d'ordre 2, de degré d = 2 et comporte lui-aussi N = 6 termes. Les modèles peuvent être décrits de manière compacte uniquement en retenant le nombre N de termes et le degré d du modèle : le seul fait de savoir si le modèle est linéaire ou non suffit alors à connaître l'ordre du modèle. Dans notre cas, nous fixons le nombre maximum N de termes permis et le degré d : ensuite le meilleur modèle est estimé avec une sélection de la structure par « rapport de réduction d'erreur » (ERR) [7]. Les modèles utilisés pour la prédiction ont donc une complexité moindre que celle de départ.

Dans chaque cas, le « meilleur modèle » correspond à l'ensemble des valeurs des paramètres θ_i obtenu par une technique de moindres carrés afin de minimiser l'erreur sur la prédiction. Une fois ces paramètres optimisés, le résidu du modèle non-linéaire est comparé à celui du modèle linéaire par un test du type Mann-Whitney. La probabilité p pour que le meilleur modèle non-linéaire soit plus performant pour prédire la dynamique sous-jacente à s(t) que le meilleur modèle linéaire est donc calculée. Ainsi, le test de non-linéarité cherche à détecter d'éventuelles relations non-linéaires entre deux échantillons décalés dans le temps, ce qui n'implique pas nécessairement la présence d'un déterminisme sous-jacent puisque la prédiction se fait sur un pas en avant, et non en prédiction libre infinie. Le modèle trouvé ne reproduit donc pas la dynamique globale du portrait de phases, mais exprime seulement un lien, linéaire ou non, entre des valeurs successives prises par une variable dynamique.

3 Dépendence à la structure du polynôme utilisé

La probabilité p d'obtenir de meilleurs résultats avec un modèle non linéaire qu'avec un modèle linéaire a été recherchée pour différentes valeurs des paramètres d et N sur le même échantillon de 2000 points d'un bruit blanc construit une seule fois.

Tab.1. Probabilité p d'obtenir une meilleure prédiction avec un modèle non linéaire qu'avec un modèle linéaire pour un bruit blanc en fonction des paramètres (d, N)qui correspondent aux valeurs maximales autorisées. Les valeurs inférieures ou égales à 0,50 sont en italique, les valeurs supérieures ou égales à 0,99 sont en gras.

d	2	3	4	5	6	7	8	9	10
N = 10	0,49	$0,\!51$	0,58	$0,\!55$	0,58	$0,\!54$	$0,\!59$	$0,\!51$	$0,\!51$
N = 20	0,44	0,47	0,53	$0,\!51$	0,53	$0,\!55$	$0,\!51$	0,50	$0,\!52$
N = 30	$0,\!51$	0,49	$0,\!55$	$0,\!51$	$0,\!52$	$0,\!58$	$0,\!54$	$0,\!52$	$0,\!60$
N = 40	$0,\!56$	0,46	$0,\!56$	0,44	0,47	$0,\!53$	0,49	0,48	$0,\!59$
N = 50	$0,\!60$	0,44	$0,\!56$	0,43	0,45	$0,\!56$	$0,\!63$	$0,\!51$	$0,\!59$
N = 60	$0,\!62$	$0,\!52$	$0,\!56$	$0,\!56$	0,50	$0,\!64$	$0,\!68$	0,72	$0,\!56$
N = 70	$0,\!67$	0,46	$0,\!51$	$0,\!51$	0,47	$0,\!68$	$0,\!66$	$0,\!72$	$0,\!86$
N = 80	0,71	0,44	$0,\!61$	$0,\!55$	0,44	$0,\!65$	0,71	$0,\!71$	$0,\!88$
N = 90	$0,\!65$	0,47	$0,\!59$	$0,\!51$	0,72	$0,\!65$	$0,\!68$	$0,\!67$	$0,\!86$
N = 100	0,71	0,47	$0,\!62$	$0,\!52$	0,77	$0,\!69$	0,76	0,72	$0,\!87$
N = 110	$0,\!80$	0,44	$0,\!59$	0,50	$0,\!78$	$0,\!67$	0,75	$0,\!75$	$0,\!89$
N = 120	$0,\!80$	0,44	$0,\!58$	0,50	$0,\!80$	$0,\!66$	0,77	$0,\!83$	$0,\!92$
N = 130	$0,\!85$	$0,\!66$	$0,\!89$	$0,\!92$	$0,\!86$	0,99	$0,\!78$	$0,\!81$	$0,\!93$
N = 140	$0,\!89$	$0,\!63$	$0,\!89$	$0,\!93$	$0,\!86$	0,99	0,76	$0,\!84$	$0,\!93$
N = 150	$0,\!90$	$0,\!65$	$0,\!90$	$0,\!95$	$0,\!88$	1,00	$0,\!79$	$0,\!86$	$0,\!95$
N = 160	$0,\!87$	$0,\!54$	$0,\!85$	$0,\!94$	$0,\!82$	0,99	0,70	$0,\!83$	$0,\!94$
N = 170	$0,\!82$	$0,\!69$	$0,\!80$	$0,\!91$	$0,\!78$	0,99	1,00	$0,\!77$	$0,\!91$
N = 180	$0,\!78$	$0,\!59$	$0,\!71$	$0,\!88$	$0,\!67$	0,99	1,00	$0,\!70$	$0,\!87$
N = 190	0,72	$0,\!50$	$0,\!63$	$0,\!83$	$0,\!60$	$0,\!98$	1,00	$0,\!62$	$0,\!85$
N = 200	$0,\!67$	0,40	$0,\!54$	0,76	0,50	$0,\!98$	1,00	$0,\!55$	$0,\!83$

Cette probabilité fluctue de façon significative avec les paramètres (d, N) retenus pour l'estimation des modèles (Tab. 1). Théoriquement, la recherche de la présence d'une composante non-linéaire au sein d'un processus aléatoire de type « bruit blanc » devrait être infructueuse. En effet, puisque nous l'avons choisi comme blanc, ce processus ne résulte pas d'un bruit sur lequel est appliqué un filtrage non linéaire. Or, le fait qu'il existe une probabilité p voisine de 0,50 indique que le choix entre un modèle linéaire et un modèle non-linéaire n'est pas évident, et que la pertinence de ces modèles est à peu près équivalente. Théoriquement, la probabilité aurait due être au voisinage de 0,5 (il ne devrait y avoir aucun avantage à utiliser un modèle non linéaire par rapport à un modèle linéaire). Aussi, il est plutôt surprenant de trouver un domaine de l'espace des paramètres pour lequel il y a un avantage évident à utiliser un modèle non linéaire ($d \approx 7$ pour N > 130) : en d'autres termes, cela revient à dire qu'une composante non linéaire a été identifiée pour ces paramètres de modélisation. Précisons que le nombre N de termes est plutôt grand, ce qui pourrait indiquer une dégradation de la qualité des modèles par diffusion du bruit numérique, notamment, celle du modèle linéaire. Ceci semble montrer que la gamme des paramètres n'est pas optimale. D'ores et déjà, nous pouvons conclure que le nombre de termes retenus, tout comme le degré du polynôme, ne devront pas être trop grands : une cinquantaine de termes avec un degré aux alentours de 3 devrait suffire.

4 Caractérisation statistique de l'outil

4.1 Cas d'un bruit blanc

Dans l'étude précédente, la fenêtre de données utilisée pour l'estimation des modèles était toujours la même. Nous souhaitons maintenant tester la robustesse de l'évaluation par rapport au choix de la fenêtre de données choisie; en d'autres termes, nous avons voulu vérifier si la probabilité p dépendait de la fenêtre choisie ou non. Les calculs suivant sont réalisés avec des modèles tels que (N = 50 et d = 3) et (N = 120 et d = 7); la comparaison entre les modèles linéaire et non linéaire est réalisée à l'aide de cinquante fenêtres différentes de 2000 points d'un même bruit blanc.



Fig. 1. Fluctuations des valeurs de p, calculées sur différents échantillons du même bruit blanc pour deux jeux de paramètres (d, N).

Pour le jeu de paramètres (d, N) = (3, 50), les probabilités p sont comprises entre 0,31 et 0,79, et pour (d, N) = (7, 120), elles fluctuent entre 0,25 et 0,94. Notons toutefois que la probabilité n'atteint jamais la valeur seuil de 0,99. Il apparaît ainsi qu'utiliser une seule valeur de la probabilité p est peu significative et qu'au moins une valeur moyenne devrait être considérée.

4.2 Cas d'une dynamique chaotique

Considérons maintenant la fonction logistique $x_{k+1} = \mu x_k (1 - x_k)$. Selon les valeurs de $\mu \in]0; 4[$, le comportement peut être périodique ou chaotique. Le paramètre μ est ici choisi égale à 3,69, valeur pour

160 E. Roulin, U. S. Freitas & C. Letellier

laquelle le comportement est chaotique. Une étude semblable à la précédente est réalisée. Il apparaît que la probabilité d'obtenir de meilleures prédictions avec un modèle non linéaire est toujours égale à 1 et ce, quelle que soit la structure du modèle choisi, et quelle que soit la fenêtre de données choisie. Ainsi, pour $\mu = 3,69$, la fonction présentant un comportement chaotique, il apparait que, sans aucune équivoque, un modèle non-linéaire est bien plus performant qu'un modèle linéaire pour prédire la dynamique décrite par la fonction logistique.

Lorsque $p \ge 0,99$, Mauricio Barahona et Chi-Sang Poon [3] proposent de procéder à un titrage du bruit, technique utilisant le principe du « virage » du titrage chimique. Il suffit d'ajouter progressivement du bruit blanc à l'échantillon de données que l'on souhaite titrer, de procéder à la détection de nonlinéarité à chaque ajout de bruit, et ce jusqu'à ce que la valeur de la probabilité p passe sous le seuil de 0,99. Au moment du « virage », la quantité de bruit ajoutée est appelée *limite de bruit* : elle correspond au rapport de la variance du bruit sur celle du signal à titrer.

La limite de bruit a ici été estimée 100 fois pour un même modèle de données (Fig. 2). À chaque nouveau titrage, la fenêtre de bruit blanc est différente. Nous retrouvons ici des résultats inhomogènes, très dépendants du bruit utilisé pour le titrage. Toutefois, cette dispersion est moindre $\left(\frac{\sigma_{LB}}{LB} = 0, 07\right)$ que dans le cas de la dépendence de p à l'échantillon de données choisi $\left(\frac{\sigma_p}{\overline{p}} = 0, 21\right)$; elle est également moins sensible à la structure du modèle. Aussi, il est fortement recommandé de réaliser plusieurs titrages pour s'assurer de la pertinence des résultats, et d'au moins en réaliser une moyenne, l'utilisation d'une valeur singulière de limite de bruit n'étant pas significative.



Fig. 2. Dispersion des valeurs de la limite de bruit calculée sur le même échantillon de la fonction logistique, lorsque les paramètres (d,N) sont fixés à (a) (3,50) et (b) (7,120). La limite de bruit est recalculée à chaque fois avec un bruit blanc différent, généré aléatoirement.

5 Dépendence au choix de l'observable

Il est connu que certaines techniques offrent des résultats qui dépendent du choix de l'observable, c'est-à-dire de la variable mesurée [8,9,10]. Ceci s'applique tout particulièrement aux systèmes d'équations différentielles où chaque variable constitue une observable possible. Prenons le cas du système de Rössler [11] :

$$\begin{cases} \dot{x} = -y - z \\ \dot{y} = x + ay \\ \dot{z} = b + z(x - c) \end{cases}$$
(3)

avec le jeu de paramètres (a, b, c) = (0.398, 2, 4) pour lequel la dynamique est chaotique. Il est alors possible d'utiliser trois observables respectivement représentées par les variables x, y et z. La notion d'observabilité repose sur les propriétés du changement de variables $\Phi_s : \mathbb{R}^3(x, y, z) \mapsto \mathbb{R}^3(s, \dot{s}, \ddot{s})$ entre l'espace des phases original $\mathbb{R}^3(x, y, z)$ et l'espace reconstruit avec les coordonnées dérivées $\mathbb{R}^3(s, \dot{s}, \ddot{s})$ induit par la variable « mesurée » s. Si Φ_s est un difféomorphisme global, la dynamique est complètement observable *via* la mesure de la variable s; c'est le cas de la variable y du système de Rössler. Malheureusement, il arrive souvent que Φ_s ne soit qu'un difféomorphisme local, c'est-à-dire qu'il existe un ensemble singulier pour lequel des états de $\mathbb{R}^3(x, y, z)$ ne soient pas distingués ou observés dans $\mathbb{R}^3(s, \dot{s}, \ddot{s})$, c'est le cas des variables x et z. Plus l'ensemble singulier est grand et proche de l'attracteur, moins la dynamique est observable [10,9]. Il est montré que les variables du système de Rössler peuvent être classées par observabilité décroissante selon $y \triangleright x \triangleright z$.

Comme dans l'exemple discuté section 3, la probabilité p dépend de la structure des modèles, c'està-dire des paramètres N et d. Contrairement au cas de la fonction logistique, il existe pour les trois variables du système de Rössler un domaine (carrés noirs) de l'espace des paramètres (Figs. 3) où les modèles ne sont pas suffisamment flexibles pour permettre au modèle non linéaire de surpasser le modèle linéaire. Ceci est logique dans la mesure où les flots sont toujours modélisés par des modèles plus compliqués que les itérations (comme les applications de premier retour). Il est donc nécessaire d'utiliser une structure suffisante. Notons que l'étendue du domaine correspondant à une mauvaise estimation des modèles dépend de la variable utilisée. Si cette surface de l'espace des paramètres (N, d) est estimée en comptant les jeux de paramètres pour lesquels le résultat obtenu est contraire à ce qui est attendu, la variable y est la plus robuste (n = 26) alors que la variable z correspond au domaine de modélisation non valide le plus important (n = 47); la variable x se situe à une position intermédiaire. Le domaine s'étend avec la dégradation de l'observabilité comme le prévoit les coefficients d'observabilité, soit $y \triangleright x \triangleright z$.



Fig. 3. Dépendence de la probabilité p à la structure des modèles (N, d) et au choix de l'observable. Un carré noir représente une probabilité p inférieure à 0,99. Cas du système de Rössler pour (a, b, c) = (0, 398, 2, 4).

En choisissant un jeu de paramètres (N = 50, d = 3) pour lequel les trois variables permettent une détection correcte de la non-linéarité, un titrage de buit est réalisé en fonction du pas de temps utilisé pour l'échantillonnage des trois séries temporelles (Fig. 4). Cinq titrages sont réalisés pour chaque valeur du pas de temps δt et pour chaque variable. Il est tout d'abord remarqué que les pas de temps pour lesquels la limite de bruit est maximale varie d'une variable à l'autre : les variables x et y ont une limite de bruit maximum pour $\delta t_{\max} \approx 1$ s alors que celle de la variable z est obtenue pour $\delta t_{\max} \approx 0, 2$ s. Le fait que le pas de temps optimal soit plus faible pour z s'explique par le fait que la nonlinéarité est confinée dans un domaine très restreint du plongement différentiel $\mathbb{R}^3(z, \dot{z}, \ddot{z})$: un pas de temps trop grand conduit nécessairement la dynamique à être perçue comme aléatoire puisque la chaîne causale est rompue. Un tel pas de temps se révèle trop petit pour les variables x et y pour lesquelles la dynamique apparaît faiblement non linéaire. Enfin, il est intéressant de noter que la valeur de la limite de bruit maximale est plus importante sur la variable z ($LB_{\max} \simeq 100$ %) que sur les deux autres variables ($LB_{\max} \simeq 60$ %) : ceci s'explique dans la mesure où la non-linéarité agit préférentiellement sur la variable z, qui la détecte donc plus fortement. Cette technique permettrait donc de détecter la variable sur laquelle la non-linéarité agit préférentiellement...

6 Conclusion

La technique de détection de non-linéarité par comparaison de modèles — linéaire versus non-linéaire — dépend en fait d'un certain nombre de paramètres dont un choix aisé est fortement recommandé. La structure des modèles utilisés ne doit être ni trop petite, ni trop grande, faute de quoi les résultats peuvent être biaisés. Si plusieurs variables sont disponibles, le résultat peut dépendre du taux d'échantillonnage



Fig. 4. Evolution des valeurs moyennes de la limite de bruit calculées pour chacune des trois variables du système de Rössler sur une fenêtre de 2000 points issus de l'intégration du système d'équation (3), pour des échantillonnages croissants. La moyenne de 5 limites de bruit est reportée avec les valeurs minimales et maximales.

utilisé et de la variable choisie : les deux peuvent s'expliquer en terme d'observabilité, les non-linéarités entraînant non seulement une perte d'observabilité mais aussi des dynamiques plus raides.

Références

- 1. U. S. FREITAS, E. ROULIN, J.-F. MUIR & C. LETELLIER, Identifying determinism underlying heart rate : the right task?, *Chaos*, **19**, 028505 (2009).
- 2. U. S. FREITAS, E. ROULIN & C. LETELLIER, Failure for distinguishing colored noise from chaos by the "Noise titration" technique, *Physical Review E.*, **79**, 035201 (2009).
- 3. C.-S. POON & M. BARAHONA, Titration of chaos with added noise, *Proceedings of the National Academy of Sciences* (USA), **98**, 7107-7112, 2001.
- M.-N. FIAMMA, C. STRAUS, S. THIBAULT, M. WYSOCKI, P. BACONNIER & T. SIMILOWSKI1, Effects of hypercapnia and hypocapnia on ventilatory variability and the chaotic dynamics of ventilatory flow in humans, *American Journal of Physiology-Regulatory Integrative and Comparative Physiology*, 292, R1985-R1993 (2007).
- L. MANGIN, M.-N. FIAMMA, C. STRAUS, J.-P. DERENNE, M. ZELTER, C. CLERICI & T. SIMILOWSKI, Source of human ventilatory chaos: Lessons from switching controlled mechanical ventilation to inspiratory pressure support in critically ill patients, *Respiratory Physiology & Neurobiology*, 161 (2), 189-196 (2008).
- M. BARAHONA & C.-S. POON, Detection of nonlinear dynamics in short noisy time series, Nature, 381, 215-217 (1996).
- L. A. AGUIRRE & S. A. BILLINGS, Improved structure selection for nonlinear models based on term clustering, *International Journal of Control*, 62 (3), 569-587 (1995).
- C. LETELLIER, J. MAQUET, L. LE SCELLER, G. GOUESBET & L. A. AGUIRRE, On the non-equivalence of observables in phase space reconstructions from recorded time series, *Journal of Physics A*, **31**, 7913-7927 (1998).
- 9. C. LETELLIER, L. A. AGUIRRE & J. MAQUET, Relation between observability and differential embeddings for nonlinear dynamics, *Physical Review E*, **71**, 066213 (2005).
- 10. C. LETELLIER & L. A. AGUIRRE, Graphical interpretation of observability in terms of feedback circuits, *Physical Review E*, **72**, 056202 (2005).
- 11. O. E. RÖSSLER, An equation for Continuous Chaos, Physics Letters A, 57 (5), 397-398 (1976).