

Construction d'un outil de Statistical Learning, basé sur un système d'Equations aux Dérivées Partielles pour l'assimilation de données d'élevage

Hélène FLOURENT, Université Bretagne Sud

Emmanuel FRENOD, Université Bretagne Sud

Vincent SINCHOLLE, NutriX

Mots-clés : Statistical Learning, EDP, Prédiction, Assimilation de données, Couplage Modèle-Données, Phénomènes biologiques

L'objectif principal de nos travaux est de construire un outil de Statistical Learning capable de prédire une sortie à partir d'une entrée concernant un animal ou un groupe d'animaux d'élevage. Nous voulons également que cet outil soit capable de réaliser de l'assimilation de données. Pour atteindre ces deux objectifs il est nécessaire de disposer d'un modèle mathématique capable de prendre en compte certains aspects de la dynamique du corps de l'animal.

L'approche consistant à modéliser avec précision certains mécanismes biologiques est une tâche nécessitant beaucoup de connaissances biologiques et menant généralement à des modèles mathématiques contenant un grand nombre d'équations [1]. Ces modèles sont utiles pour décrire et comprendre des processus biologiques spécifiques mais leur complexité limite leur flexibilité et leur adaptation aux données et aux processus biologiques étudiés. D'autre part, les outils de Machine Learning classiques de type réseaux de neurones sont souvent limités pour prendre en compte l'existence de nombreux phénomènes biologiques sous-jacents. En effet ces outils permettent d'intégrer facilement des variables d'entrée et de sortie mais ils sont basés sur des modèles mathématiques relativement simples dans lesquels il est difficile d'intégrer des phénomènes complexes (accumulation, saturation, retard, etc.)[2].

L'objectif de notre étude est de construire un outil capable d'allier précision, parcimonie et flexibilité. Pour cela nous avons choisi d'explorer une approche de type Couplage Modèle-Données [3], couplant l'utilisation de connaissances pour construire un modèle mathématique parcimonieux et l'utilisation de données pour ajuster les paramètres de ce modèle.

Nous avons donc construit un modèle mathématique basé sur un système d'Equations aux Dérivées Partielles (EDP) qui embarquent l'expression mathématiques de phénomènes comme la diffusion, la convection, la fixation, etc. Chacune de ces EDP contient des paramètres associés à des facteurs biologiques (vitesse de convection, seuils de saturation, coefficient de retard, etc.) pouvant être appris à partir de données et conférant donc à notre outil une capacité d'apprentissage.

Afin de tester la capacité d'apprentissage de notre outil, des tests par simulations ont été réalisés. Une première application de l'outil sur des données collectées en élevage nous a ensuite permis de tester les capacités du modèle sur des données réelles.

Références

- [1] D. BASTIANELLI, D. SAUVANT AND A. RERAT, *Mathematical modeling of digestion and nutrient absorption in pigs.*, Journal of animal science, 1996.
- [2] M. T. GORCZYCA, H. F. M. MILAN, A. S. C. MAIA AND K. G. GEBREMEDHIN, *Machine learning algorithms to predict core, skin, and hair-coat temperatures of piglets*, Computers and Electronics in Agriculture, 2018.
- [3] W. SACKS, D. SCHIMEL, R. K. MONSON AND B. H. BRASWELL, *Model-data synthesis of diurnal and seasonal CO₂ fluxes at Niwot Ridge, Colorado*, Global Change Biology, 2006.

Hélène FLOURENT, Université Bretagne Sud, Laboratoire de Mathématiques de Bretagne Atlantique,UMR CNRS 6205, Campus de Tohannic, Vannes, France
helene.flourent@univ-ubs.fr

Emmanuel FRENOD, Université Bretagne Sud, Laboratoire de Mathématiques de Bretagne Atlantique,UMR CNRS 6205, Campus de Tohannic, Vannes, France
emmanuel.frenod@univ-ubs.fr

Vincent SINCHOLLE, NutriX, France