

Qu'est ce qui fait nager les tortues? :

Un projet en fouille de données appliquées à l'analyse de trajectoires de tortues marines et de données océanographiques.

1. CONTEXTE

Toutes les espèces de tortues marines sont actuellement en danger, voire menacées d'extinction pour certaines populations. La mise en place de mesures de protection appropriées est difficile, notamment pour les tortues au stade « juvénile pélagique ». En effet, on ne dispose que de très peu d'informations sur la façon dont les tortues se dispersent au large entre le moment où elles partent en mer juste après leur naissance et le moment où, adultes, elles reviennent pondre à terre pour la première fois. Cette période, souvent appelée « les années perdues » dure de quelques années à quelques décennies selon les espèces.

Néanmoins, des chercheurs de la NOAA (National Oceanic and Atmospheric Administration, USA) ont accumulé un impressionnant jeu de données de trajectoires de tortues caouanne juvéniles dans l'océan Pacifique Nord. Ils ont en effet suivi, entre 1997 et 2013 plus de 200 tortues caouannes juvéniles équipées de balises ARGOS. Grâce à un accord de collaboration avec ces chercheurs nous avons désormais accès à ce jeu de données. Il contient au total plus de 230 années de suivi, soit plus de 85 000 positions et vecteurs vitesse journaliers. Un premier travail de stage a permis d'extraire les valeurs de différentes variables environnementales en chacun des points/dates d'observation, notamment :

- la température de l'eau (SST) et la production primaire (NPP, un proxy pour la quantité de nourriture disponible)
- les gradients méridiens et zonaux de la SST et de la NPP
- la vitesse du courant (méridienne et zonale)
- Les 3 composantes du vecteur champ magnétique terrestre
- La longueur du jour

2. OBJECTIF DU TRAVAIL

L'objectif du travail proposé est d'analyser ce jeu de données très complet pour déterminer dans quel mesure il est possible de prévoir les mouvements des tortues juvéniles à partir de prévisions de l'état de l'océan. L'objectif ultime est de disposer d'un modèle de prévision des zones où la probabilité de présence de tortues est forte afin de limiter les activités de pêche dans ces zones et minimiser ainsi les risques de captures accidentelles de tortues marines.

Il s'agira donc de tester si, avec une approche de type machine learning, on peut prédire (au moins en partie) la vitesse de nage des tortues juvéniles à partir des variables environnementales figurant dans ce jeu de données. La réponse est très certainement « oui » pour certaines variables telles que la température puisqu'on sait que les tortues marines sont

des animaux ectothermes qui doivent se déplacer pour maintenir leur température corporelle dans une gamme de températures acceptables. Cependant, il sera intéressant (par exemple) de déterminer si le refroidissement de l'eau est le seul prédicteur utile de l'amorce de la migration vers le sud au début de l'automne ou si la diminution de la longueur du jour est susceptible de jouer un rôle?

Pour atteindre cet objectif, une phase préliminaire d'analyse exploratoire des données incluant les démarches classiques univariées, bivariées et multivariées, devra être menée. Ensuite, la question de la modélisation de la vitesse de nage pourra d'abord être abordée par des méthodes linéaires (régression linéaire multiple), intégrant ou pas des techniques de sélection de variables (pénalisation LASSO, elastic net). Enfin, les méthodes non linéaires du domaine du machine learning (réseaux de neurones, SVM, random forest) pourront être mises en œuvre pour traiter le problème de prédiction. Une démarche de validation croisée sera à envisager pour évaluer la qualité des modèles de prédiction.

Logiciels utilisés : R ou Python. Une attention particulière devra être apportée à l'écriture des codes en vue d'une utilisation ultérieure à l'issue du projet.

Encadrement :

Philippe Gaspar (Mercator-Océan) : pgaspar@mercator-ocean.fr

Sébastien Dejean (Institut de Mathématiques de toulouse) : sebastien.dejean@math.univ-toulouse.fr