**Partie 1 : Estimation de la Probabilité de Transmission à l'Aide du Modèle Binomial**

**Contexte et Objectif**

Dans le cadre de notre étude sur la transmission d'une lignée virale affectant les petits ruminants, nous avons entrepris d'estimer la probabilité de transmission de cette maladie dans un milieu expérimental. Pour ce faire, nous avons adopté une approche statistique basée sur le modèle binomial, considérant la nature discrète des résultats d'expérimentation (transmission réussie ou non).

**Méthodologie**

Notre analyse repose sur un ensemble de 16 expérimentations, chacune impliquant 6 animaux. L'objectif était de déterminer le nombre de cas positifs (transmissions réussies) après chaque expérimentation. Pour modéliser cette situation, nous avons utilisé un modèle binomial, où la probabilité de succès varie en fonction de la durée de chaque expérimentation selon la formule : , avec représentant une probabilité fixe de transmission pour toutes les expérimentations et représente la durée de l’expérimentation. Afin d'inférer la valeur de , nous avons opté pour une approche bayésienne, utilisant une prior Uniforme pour , ce qui reflète notre absence d’informations sur la valeur de avant l'analyse. La vraisemblance des données observées a été modélisée par une distribution Binomiale, exprimée comme suit : , où représente le nombre d'animaux par expérimentation et le nombre de cas positifs observés.

**Analyse Statistique**

Pour l'échantillonnage de la distribution postérieure de notre modèle, nous avons utilisé les chaines de Markov Monte Carlo (MCMC) avec l’échantillonneur NUTS (No-U-Turn Sampler), le sampler par défaut de la bibliothèque PyMC sous Python. La configuration de notre échantillonnage comprenait 4 chaînes, avec 2000 itérations chacune, précédées de 1000 itérations de tuning pour affiner les paramètres de l'algorithme. Un paramètre target\_accept de 0.95 a été fixé pour minimiser la divergence des chaînes, assurant ainsi la fiabilité de nos estimations. Pour étudier la qualité et la convergence des chaines nous avons utilisé le diagnostic R-hat de Gelman-Rubin et la mesure Effective Sample Size (ESS).

**Résultats et Interprétation**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Mean | Sd | Hdi\_3% | Hdi\_97% | Mcse\_mean | Mcse\_sd | Ess\_bulk | Ess\_tail | R\_hat |
|  | 0.000193 | 0.000042 | 0.000113 | 0.000268 | 0.000001 | 0.000001 | 1684.265199 | 1412.719496 | 1.001427 |

**A screenshot of a graph

Description automatically generated**

Figure : Distribution et trace du paramètre estimé p

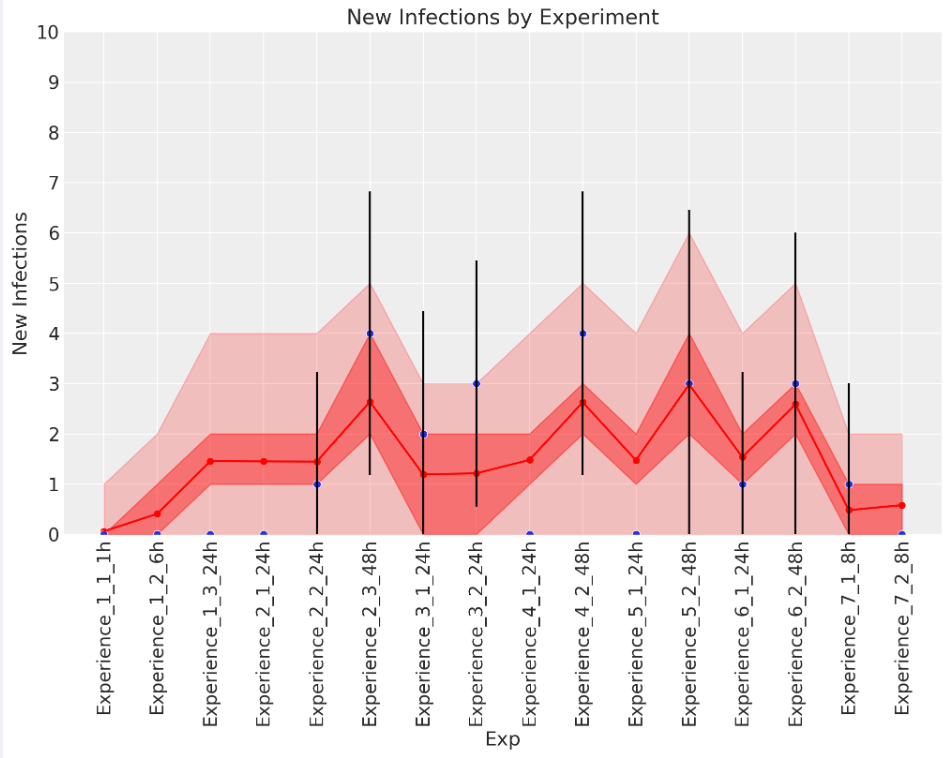


Figure : Comparaison entre les observations réelles en points bleus et les estimations en points rouges

**Partie 2 : Modélisation de la Transmission en Fonction de la Distance et du Temps d'Exposition (fixes)**

**Méthodologie**

Afin d'affiner notre compréhension de la dynamique de transmission de la lignée virale IV du virus PPRV, nous avons développé une approche prenant en compte la distance entre les paires d’animaux (animal susceptible + seed) ainsi que le temps d'exposition à cette distance. Cette méthode nous a permis de modéliser la probabilité de transmission en fonction de variables spatiales et temporelles spécifiques à chaque animal. Pour chaque animal, nous avons calculé le temps cumulé d'exposition en fonction de trois distances prédéfinies par rapport à l'animal infecté : 0.3 m, 1 m, et 2 m. Ces distances ont été choisies pour refléter des scénarios d'exposition variés au sein du milieu expérimental. La probabilité individuelle de transmission a été modélisée comme suit : , où représente le temps d'exposition cumulé pour l'animal considéré.

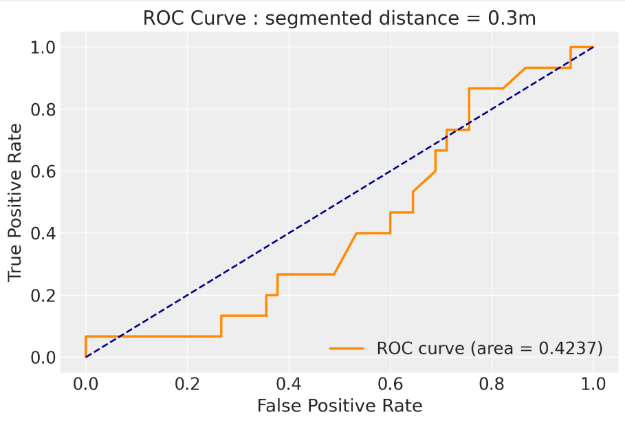
**Modélisation Statistique**

Pour chaque distance, un modèle distinct a été construit et exécuté, utilisant la distribution Bernoulli pour modéliser la transmission (infecté / non infecté) comme une variable binaire observée. La configuration de l'échantillonnage MCMC pour l'inférence a été la suivante : 10 chaînes, 10 000 itérations par chaîne, avec 1 000 itérations de tuning préalables, et un paramètre target\_accept de 0.96 pour optimiser la convergence des chaînes. Cette configuration a été appliquée uniformément à tous les modèles pour garantir la cohérence de l'analyse.

**Résultats et Interprétation**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Distance | Mean | Sd | Hdi\_3% | Hdi\_97% | Mcse\_mean | Mcse\_sd | Ess\_bulk | Ess\_tail | R\_hat |
|  | 0.3 m | 0.179649 | 0.040335 | 0.105006 | 0.255019 | 0.000221 | 0.000156 | 32991.135315 | 30256.250373 | 1.00021 |
| 1 m | 0.004886 | 0.001188 | 0.002793 | 0.007166 | 0.000007 | 0.000005 | 30647.167723 | 27226.184696 | 1.000221 |
| 2 m | 0.001951 | 0.000465 | 0.001112 | 0.002801 | 0.000003 | 0.000002 | 20556.141899 | 16071.877778 | 1.000309 |

|  |  |  |  |
| --- | --- | --- | --- |
| Distance | p\_waic | p\_loo | mean\_AUC |
| 0.3 m |  |  |  |
| 1 m |  |  |  |
| 2 m |  |  |  |

A graph of a curve

Description automatically generated

A graph of a curve

Description automatically generatedA graph of a curve

Description automatically generated

**Partie 3 : Enveloppes de distances**

La troisième approche utilise des enregistrements de distances par minute, permettant une analyse plus granulaire et dynamique des interactions entre les animaux. Contrairement à la deuxième approche, qui se base sur des distances fixes (1 m, 1.5 m et 2 m) et le temps cumulé d'exposition, cette méthode capture les variations minutieuses des distances entre les animaux au fil du temps. Au lieu de considérer des distances prédéfinies, la troisième approche identifie des groupes où la distance entre les animaux se stabilise. Pour chaque groupe, la distance médiane et le temps total de stabilisation sont calculés. Cette stratégie permet de détecter des patterns de contacts et de mouvements plus complexes et réalistes, reflétant mieux les interactions naturelles entre les animaux.

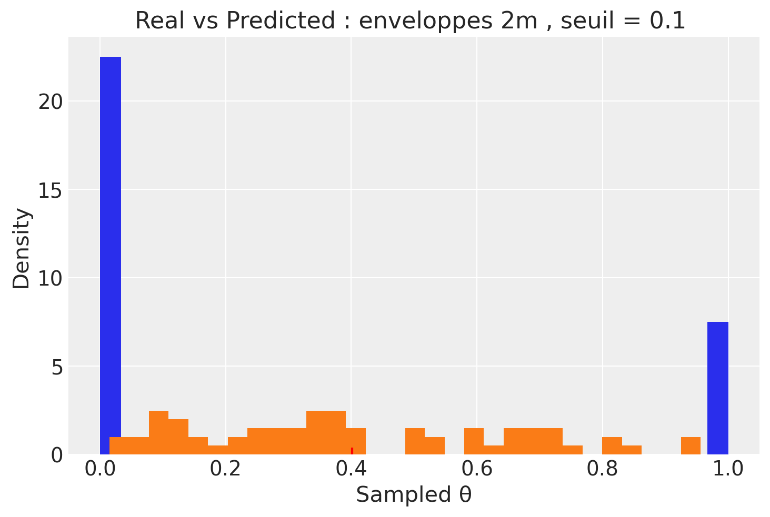
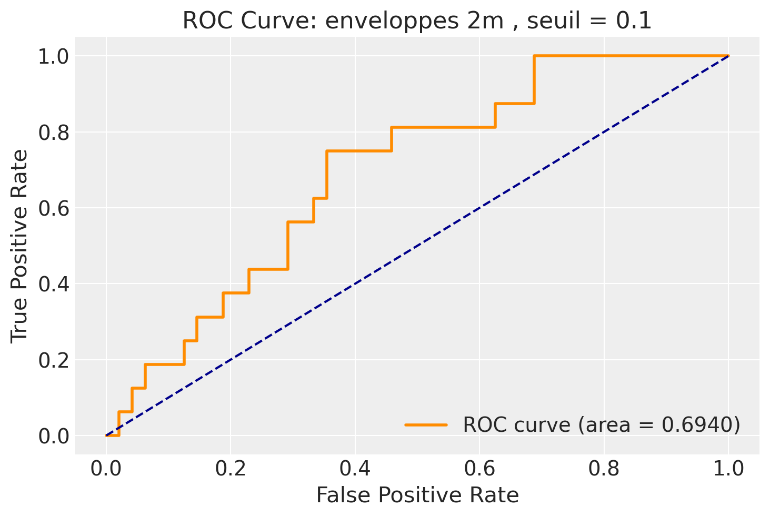
**Modélisation de la Probabilité de Transmission :**

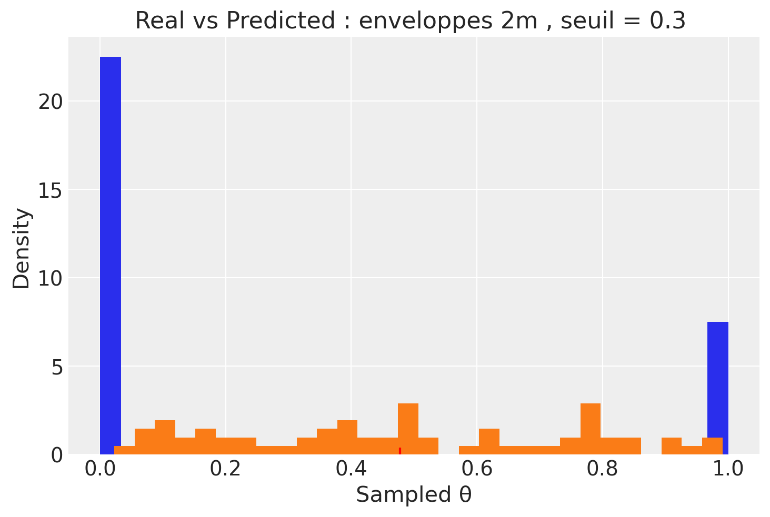
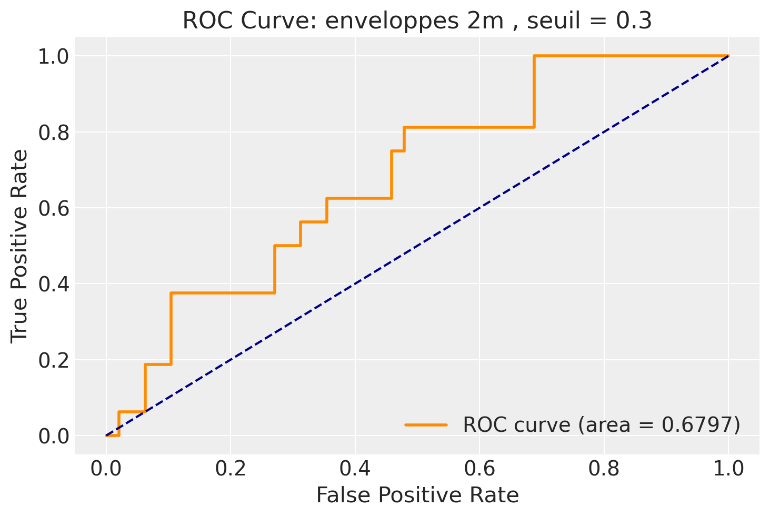
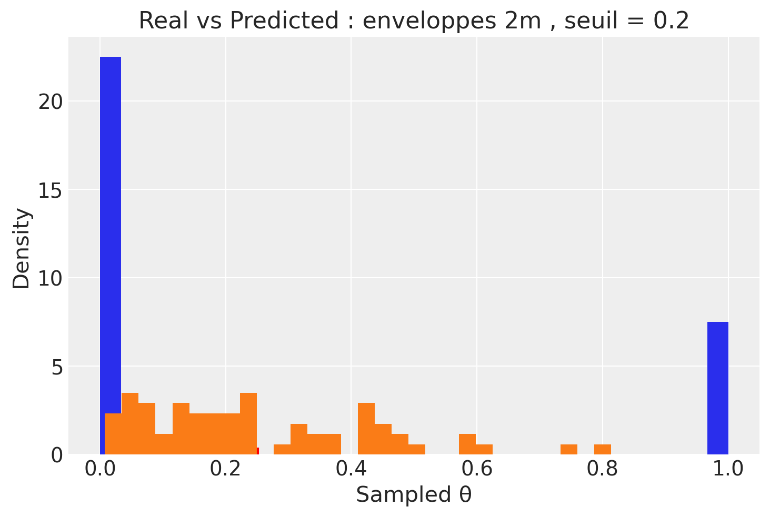
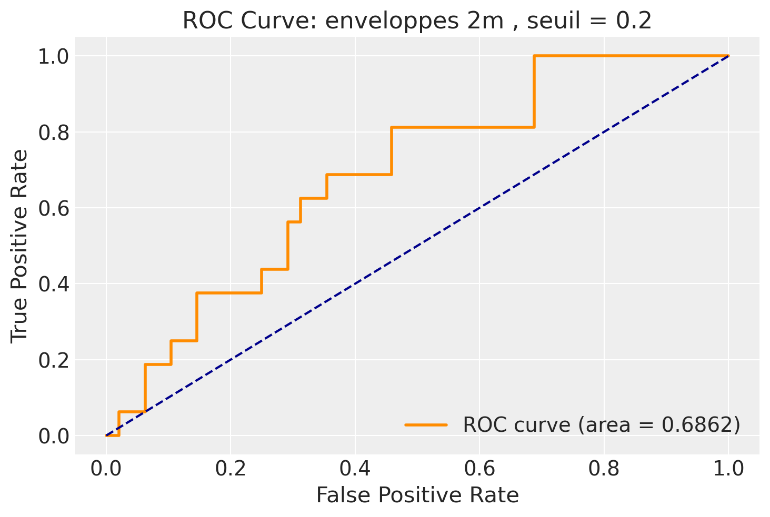
La formule de probabilité de transmission dans la troisième approche : , introduit une composante exponentielle qui prend en compte à la fois la distance et le temps de manière plus sophistiquée. Cette formule permet de modéliser l'effet de la distance sur la probabilité de transmission de manière non linéaire, contrairement à la formule linéaire simplifiée utilisée dans la deuxième approche.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Distance | Seuil | Paramètre | Mean | 94% HDI |
| 1 m | 0.1 |  | 0.0083 | [0.0045 – 0.012] |
|  | 0.47 | [0.39 – 0.56] |
| 0.2 |  | 0.0088 | [0.0048 – 0.013] |
|  | 0.48 | [0.39 – 0.57] |
| 0.3 |  | 0.0092 | [0.005 – 0.014] |
|  | 0.48 | [0.39 – 0.57] |
| 1.5 m | 0.1 |  | 0.0043 | [0.0024 – 0.0063] |
|  | 0.34 | [0.28 – 0.41] |
| 0.2 |  | 0.0043 | [0.0024 – 0.0064] |
|  | 0.34 | [0.28 – 0.41] |
| 0.3 |  | 0.0053 | [0.003 –0.0078] |
|  | 0.34 | [0.28 – 0.41] |
| 2 m | 0.1 |  | 0.0034 | [0.0019 – 0.0051] |
|  | 0.37 | [0.3 – 0.44] |
| 0.2 |  | 0.0035 | [0.002 – 0.0052] |
|  | 0.37 | [0.3 – 0.44] |
| 0.3 |  | 0.0032 | [0.0018 – 0.0046] |
|  | 0.25 | [0.2 – 0.3] |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Distance | Seuil | mean\_AUC | 95% CI | p\_waic | p\_loo |
| 1 m | 0.1 |  |  |  |  |
| 0.2 |  |  |  |  |
| 0.3 |  |  |  |  |
| 1.5 m | 0.1 |  |  |  |  |
| 0.2 |  |  |  |  |
| 0.3 |  |  |  |  |
| 2 m | 0.1 | 0.6898 | (0.6868, 0.6921) | 0.96 | 0.96 |
| 0.2 | 0.6825 | NA | 0.99 | 0.99 |
| 0.3 | 0.6774 | (0.6699, 0.6855) | 0.98 | 0.98 |

Distance = 2 m





Distance = 1 m