

Rapport 28/10/2022

Contents

1	Sans perturbations	2
1.1	Estimation de l'effet de la profondeur	2
1.2	Recherche de la mémoire optimale	2
1.3	Régression bi-exponentielle	4
2	Estimation de l'effet de l'exposition	5

Nous souhaitons observer chez des narvals l'effet de l'exposition à des perturbations humaines sur leur capacité à se nourrir. Lorsqu'elles se nourrissent, ces baleines émettent des sons spécifiques (buzz). Nous nous intéresserons donc à la fréquence d'émission de ces buzz (buzz/min), ce qui peut être modélisé par un modèle de Poisson.

1 Sans perturbations

1.1 Estimation de l'effet de la profondeur

Pour se nourrir, les narvals doivent plonger profondément alors que le reste du temps, ils restent "proche" de la surface. Il faut donc inclure au modèle la profondeur à laquelle se trouvent les baleines quand elles émettent ou non des buzz. La relation entre l'émission de buzz et la profondeur n'étant pas linéaire, la profondeur a été remplacée par une variable explicative la décrivant par un polynôme de degré 3.

1.2 Recherche de la mémoire optimale

L'émission d'un buzz à un instant t est corrélé à l'émission ou non de buzz aux instants précédents t . Cet effet mémoire doit donc être ajouté au modèle et pour cela nous devons déterminer la mémoire maximum qu'il faut autoriser au modèle.

Afin de trouver la mémoire optimale, nous avons utilisé la démarche suivante :

1. $from.ml = 1$; $to.ml = N$
2. Tant que $(to.ml - from.ml > 2)$:
 1. Pour $ml_i^k = from.ml + (i - 1) * \lfloor \frac{to.ml - from.ml}{M-1} \rfloor$, $i = 1...M$:
 1. ajustement d'un modèle de Poisson avec ml_i^k éléments mémoire
 2. calcul du BIC
 2. $i_{opt} = \operatorname{argmin} BIC_i$
 3. $ml_{opt} = ml_{i_{opt}}^k$
 4. $from.ml = ml_{i_{opt}-1}^k$; $to.ml = ml_{i_{opt}+1}^k$

1.2.1 Sans effets aléatoires

Dans un premier temps, nous n'avons pas inclus d'effets aléatoires sur les individus pour estimer la mémoire optimale :

$$Buzz \sim Ind + \text{spline}(Depth) + Lag_1 + \dots + Lag_N$$

Nous avons fixé $N = 300$ et $M = 10$.

Nous obtenons une mémoire optimale de 60 secondes. La figure 1 permet de voir que cet optimal semble bien correspondre à un minimum global.

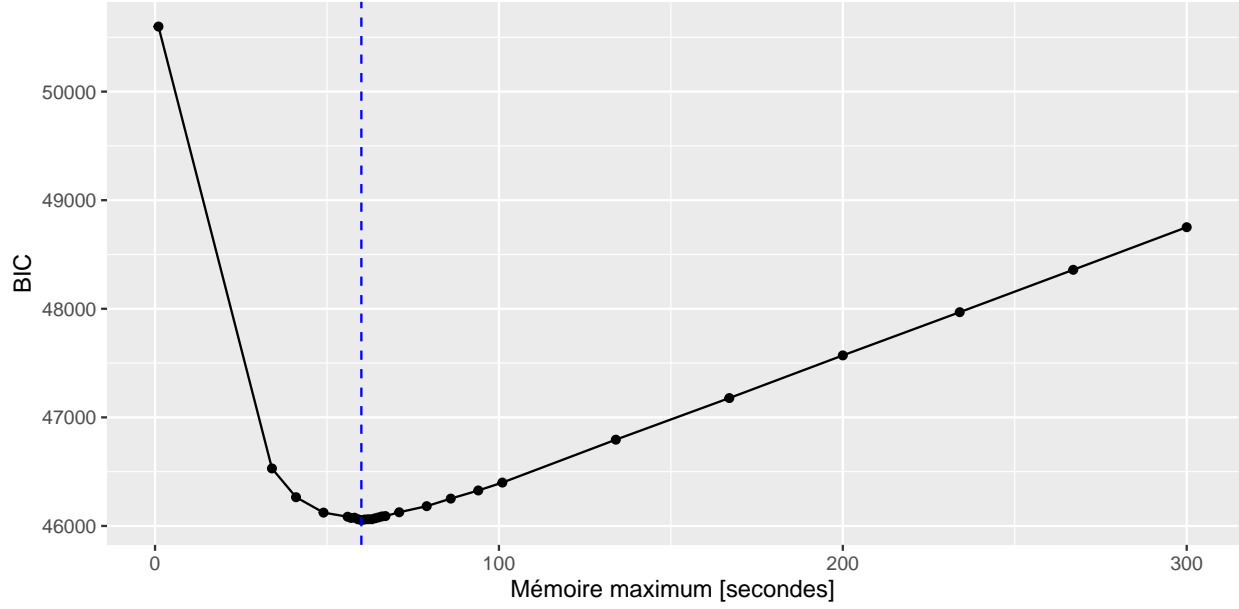


Figure 1: BIC en fonction de la mémoire maximum

1.2.2 Avec effets aléatoires

Nous avons plusieurs observations par individu, ce qui implique qu'elles ne sont pas indépendantes. Pour prendre en compte cette dépendance, nous avons tenté d'inclure un effet aléatoire sur les individus :

$$Buzz \sim (1|Ind) + \text{spline}(Depth) + Lag_1 + \dots + Lag_N$$

Compte tenu du résultat obtenu précédemment, nous avons tout de suite restreint l'intervalle de recherche initial à $[30, 90]$.

Le temps d'ajustement des modèles mixtes (environ 1h30) ne nous a pas permis d'obtenir de résultats définitifs pour le moment, la tâche tourne encore sur le cluster du Gricad.

La figure 2 présente les résultats temporaires. La courbe semble pour le moment similaire à celle obtenue sans l'ajout d'effets aléatoires.

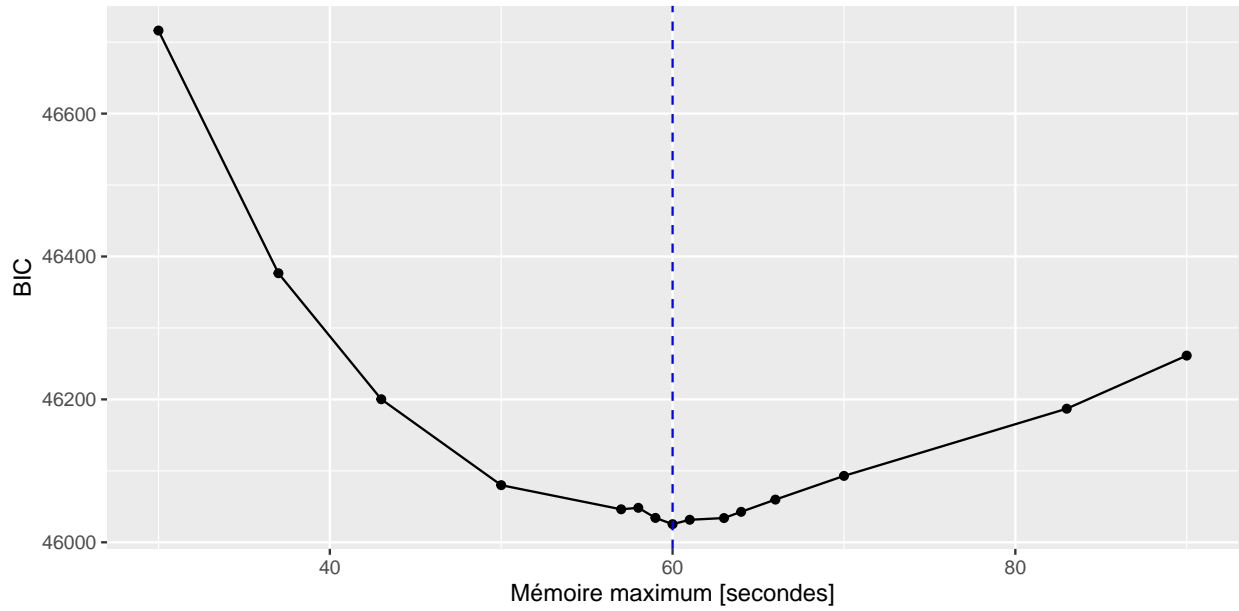


Figure 2: BIC en fonction de la mémoire maximum - modèle mixte

1.3 Régression bi-exponentielle

Nous avons vu que les modèles mixtes sont bien plus longs à ajuster que les modèles “classiques”, aussi, afin de réduire le nombre de variables, nous appliquons une régression bi-exponentielle aux coefficients auto-régressifs obtenus précédemment. Ainsi nous passons de 60 à 4 variables.

La figure 3 montre les composantes de la mémoire ajustées pour un décalage maximum de 60. Les points sont les contributions individuelles de chaque décalage, la courbe correspond à la régression bi-exponentielle.

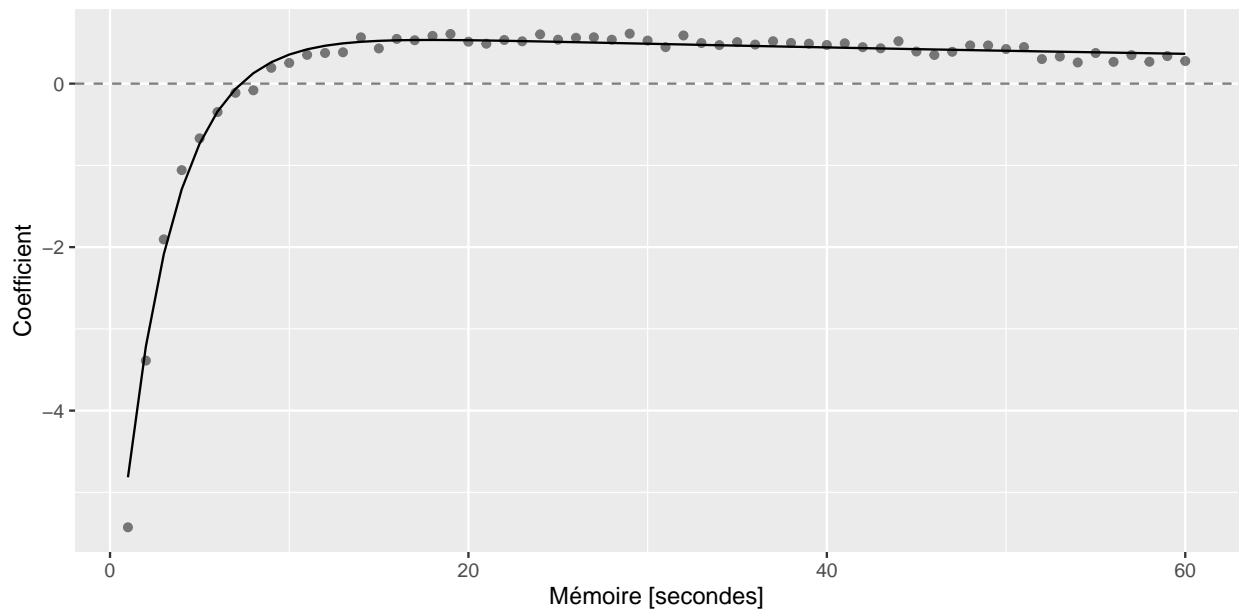


Figure 3: Régression bi-exponentielle des coefficients auto-régressifs

Les coefficients obtenus suite à la régression bi-exponentielle sont présentés dans la table 1.

Table 1: Coefficients obtenus par régression bi-exponentielle

b1	b2	b3	b4
-7.739335	0.3483068	0.653957	0.0097382

2 Estimation de l'effet de l'exposition

L'exposition aux perturbations est exprimée par $1/dist$ où $dist$ est la distance en kilomètres séparant l'animal du bateau dont émane la perturbation.

Nous avons ajusté le modèle mixte suivant :

$$Buzz \sim (1|Ind) + offset(ARDepth) + spline(Expo)$$

Table 2: Coefficients du modèle mixte incluant l'exposition

effect	group	term	estimate	std.error	statistic	p.value
fixed	NA	(Intercept)	-4.5901334	0.0352823	-130.0972	0
fixed	NA	ns(X, knots = quantile(data\$X[data\$X > 0], 1:2/3))1	-1.1481117	0.0021543	-532.9509	0
fixed	NA	ns(X, knots = quantile(data\$X[data\$X > 0], 1:2/3))2	-58.8986239	0.0543053	-1084.5839	0
fixed	NA	ns(X, knots = quantile(data\$X[data\$X > 0], 1:2/3))3	-112.8364412	0.1069249	-1055.2866	0
ran_pars	Ind	sd_(Intercept)	0.0864238	NA	NA	NA

Table 3: Statistiques du modèle mixte incluant l'exposition

sigma	logLik	AIC	BIC	deviance	df.residual
1	-34391987382	68783974775	68783974838	51672378818	2367453

Ci-dessous, plusieurs graphiques utiles pour valider visuellement le modèle. Nous n'avons pas eu trop le temps de les discuter, mais ils semblent encourageants. Nous avons tout de même remarqué une valeur franchement aberrante.

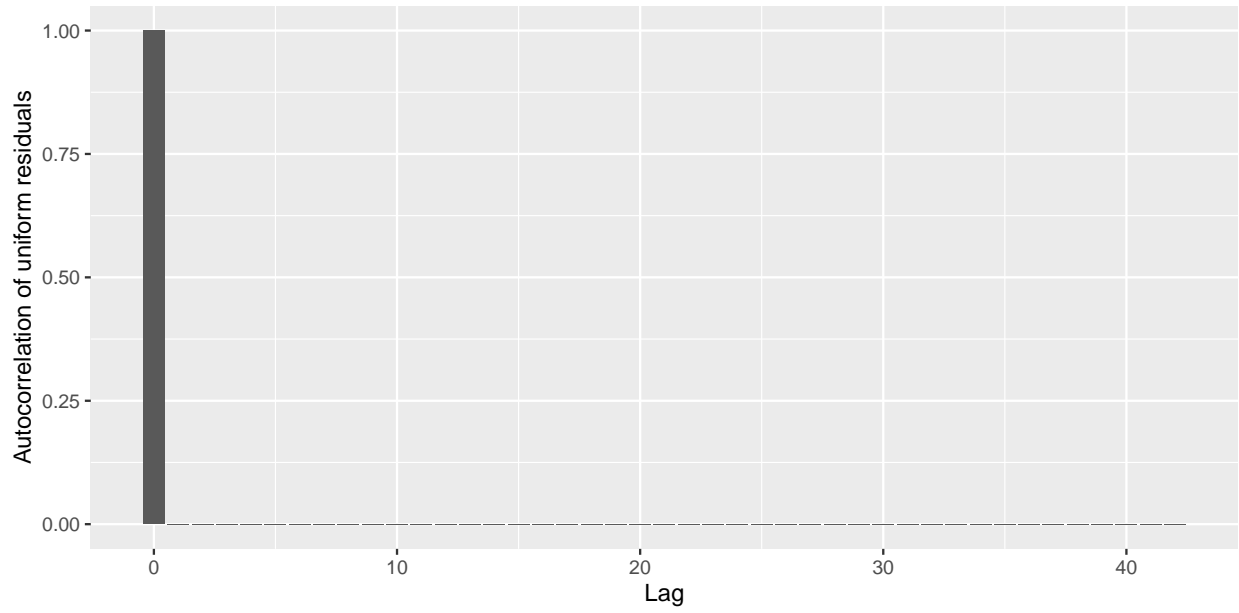


Figure 4: Auto-corrélation des résidues

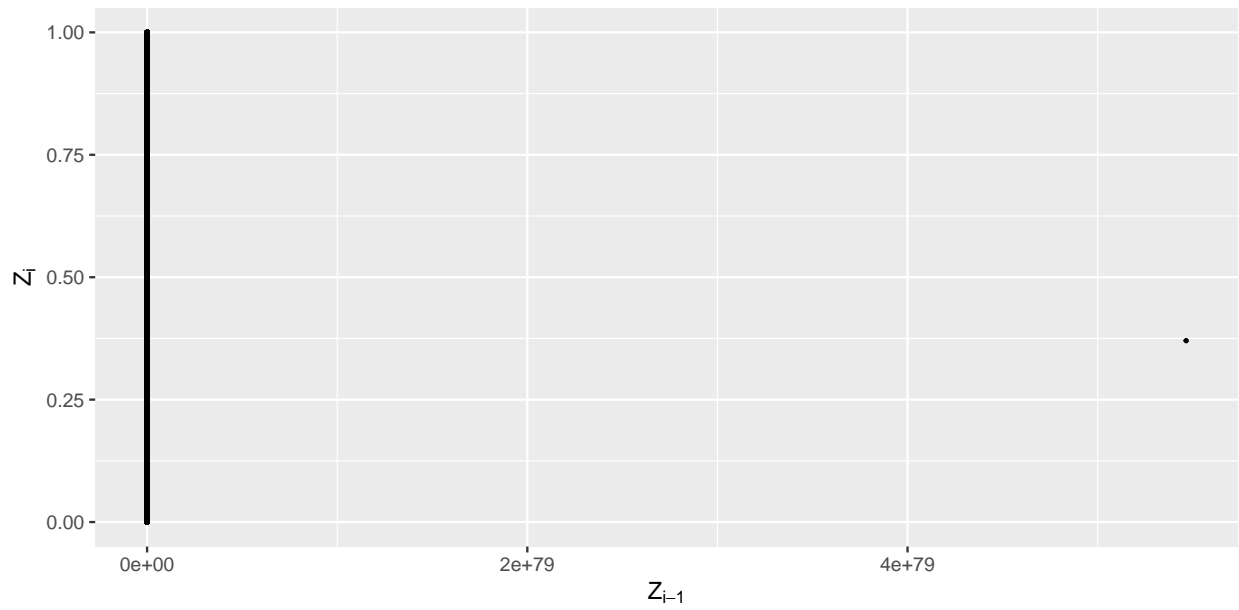


Figure 5: Résidue selon le précédent. Avec la valeur aberrante

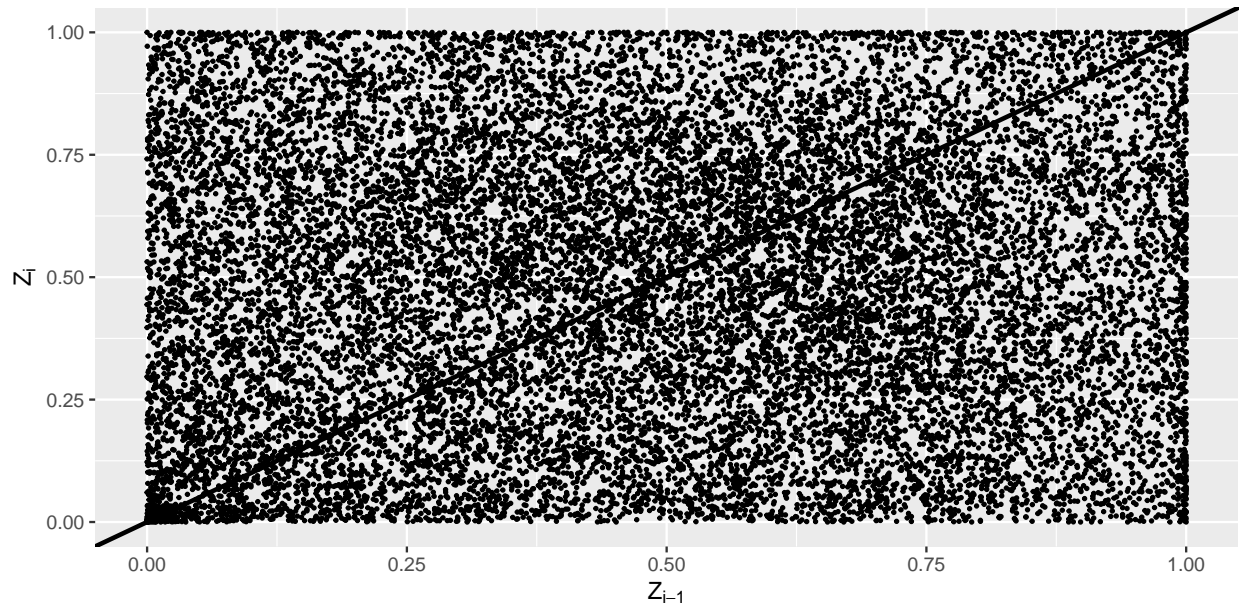


Figure 6: Résidue selon le précédent

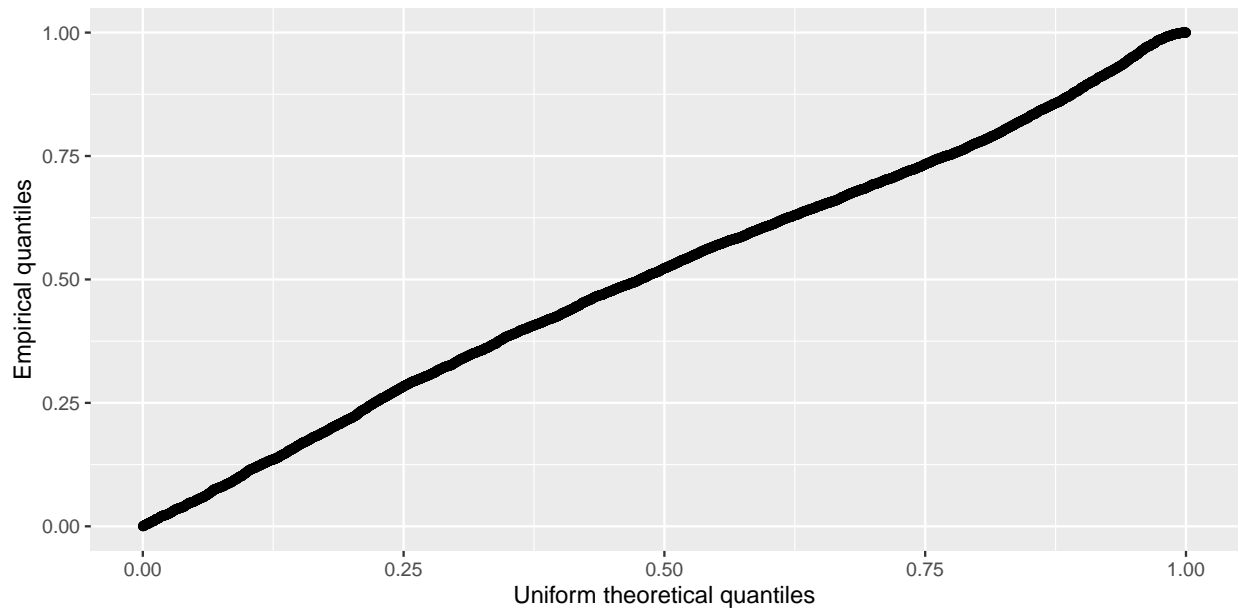


Figure 7: Quantiles-Quantiles