L'estimation des mortalités : éléments clés pour leur réalisation...et leur bon usage.

Aurélien Besnard

Ecole Pratique des Hautes Etudes
UMR5175 Centre d'Ecologie Fonctionnelle et Evolutive
Labex CEMEB



















Principes généraux de la mesure des impacts

- > Prospections répétées sous les éoliennes
- > Collecte de cadavres

...mais comment passer d'un nombre de cadavres trouvés à un nombre total de mortalités?





Principes généraux de la mesure des impacts

- > Prospections répétées sous les éoliennes
- > Collecte de cadavres

...mais comment passer d'un nombre de cadavres trouvés à un nombre total de mortalités?

- (i) La prospection n'est pas exhaustive (on ne passe pas en continu) il s'agît d'un échantillonnage qui nécessite de produire des estimations avec un degré de confiance déterminé (processus aléatoire)
- (ii) Il existe de nombreux biais liés aux prospections nécessite de corriger ces biais



Deux composantes classiques à gérer pour l'estimations des mortalités : Corriger les biais, caractériser la précision des estimateurs





Les biais liés à la persistance des cadavres

Les charognards consomment les cadavres, les cadavres se dégradent

ils ne sont donc plus **observables**

Conséquence : le nombre de cadavres « disponible » est une sous-estimation du nombre réel

Les biais liés à la persistance des cadavres

Les charognards consomment les cadavres, les cadavres se dégradent

ils ne sont donc plus **observables**

Conséquence : le nombre de cadavres « disponible » est une sous-estimation du nombre réel

Ce biais est **très hétérogène** dans l'espace et le temps, il dépend:

- de l'abondance et du type de charognards
- de la saison (abondance de proies alternatives, période de reproduction, etc)
- de la taille des espèces
- de « l'âge » des cadavres
- de la localisation même du cadavre (« micro-habitat »)

Les biais liés à la persistance des cadavres

Les charognards consomment les cadavres, les cadavres se dégradent

ils ne sont donc plus **observables**

Conséquence : le nombre de cadavres « disponible » est une sous-estimation du nombre réel

Ce biais est **très hétérogène** dans l'espace et le temps, il dépend:

- de l'abondance et du type de charognards
- de la saison (abondance de proies alternatives, période de reproduction, etc)
- de la taille des espèces
- de « l'âge » des cadavres
- de la localisation même du cadavre (« micro-habitat »)

Recommandation : **estimer localement la durée de persistance** des cadavres avec des expérimentations spécifiques (dépose de cadavres) mais :

- Des problèmes de **crédibilité** (où sont placés les cadavres, odeur de l'observateur ou du cadavre, densité de cadavres mis en place)
- Ce sont des estimations qui doivent être associées à un intervalle de confiance qui dépend du nombre de cadavres (taille de l'échantillon)

Les biais de détection des cadavres (toujours présents)

Tous les cadavres présents **ne sont pas trouvés** par les observateurs.

Conséquence: le nombre de cadavres trouvés est une sous-estimation du nombre réel.





Les biais de détection des cadavres (toujours présents)

Tous les cadavres présents **ne sont pas trouvés** par les observateurs.

Conséquence: le nombre de cadavres trouvés est une sous-estimation du nombre réel.

Ce biais est très hétérogène dans l'espace et le temps, il dépend:

- du milieu (fermé, ouvert...et cela peut évoluer au cours des saisons)
- des observateurs (expérimentés ou pas, selon méthode de prospection, etc)
- de l'état de l'observateur (fatigue, concentration, etc)
- de la taille des espèces





Les biais de détection des cadavres (toujours présents)

Tous les cadavres présents ne sont pas trouvés par les observateurs

Conséquence: le nombre de cadavres trouvés est une sous-estimation du nombre réel

Ce biais est très hétérogène dans l'espace et le temps, il dépend:

- du milieu (fermé, ouvert...et cela peut évoluer au cours des saisons)
- des observateurs (expérimentés ou pas, selon méthode de prospection, etc)
- de l'état de l'observateur (fatigue, concentration, etc)
- de la taille des espèces

Recommandation : **estimer localement la probabilité de détection** des cadavres avec des expérimentations spécifiques (dépose de cadavres) mais :

- Des problèmes de **crédibilité** (où sont placés les cadavres, l'observateur sait qu'il y en a, etc)
- Ce sont des estimations qui doivent être associées à un intervalle de confiance qui dépend du nombre de cadavres (taille de l'échantillon)





Les biais de prospection surfacique

Toute la surface dans laquelle peuvent tomber les cadavres n'est pas couverte.

Conséquence: le nombre de cadavres trouvés est une sous-estimation du nombre réel.





Les biais de prospection surfacique

Toute la surface dans laquelle peuvent tomber les cadavres n'est pas couverte.

Conséquence: le nombre de cadavres trouvés est une sous-estimation du nombre réel.

Un biais très hétérogène dans l'espace et le temps:

- entre milieux (fermés, ouverts...et il peut évoluer selon les saisons)
- entre observateurs (selon méthode de prospection, etc)





Les biais de prospection surfacique

Toute la surface dans laquelle peuvent tomber les cadavres n'est pas couverte.

Conséquence: le nombre de cadavres trouvés est une sous-estimation du nombre réel.

Un biais très hétérogène dans l'espace et le temps:

- entre milieux (fermés, ouverts...et il peut évoluer selon les saisons)
- entre observateurs (selon méthode de prospection, etc)

Recommandation : corriger pour les surfaces vraiment prospectées mais :

- le nombre de cadavres décroit avec la distance à l'éolienne
- la surface de l'anneau augmente linéairement avec la distance
- la densité de cadavres baisse donc fortement avec la distance
- la surface réellement prospectée peut varier selon la distance



Possible de corriger en calculant sur des anneaux...mais pose des questions d'estimation (faibles tailles d'échantillon dans chaque anneau, donc intervalles de confiance larges)





Le « Cripping loss »

Terme dérivé de la littérature scientifique en modélisation des impacts de la chasse.

Il s'agit de tous les individus qui meurent en dehors de la zone prospectée

- soit à proximité
- soit dans les heures ou jours qui suivent, donc à de grandes distances

Conséquence : le nombre de cadavres trouvés est une sous-estimation du nombre réel





Le « Cripping loss »

Terme dérivé de la littérature scientifique en modélisation des impacts de la chasse.

Il s'agit de tous les individus qui meurent en dehors de la zone prospectée

- soit à proximité
- soit dans les heures ou jours qui suivent, donc à de grandes distances

Conséquence : le nombre de cadavres trouvés est une sous-estimation du nombre réel

Recommandation : pas de solution, aucune étude disponible...





Bilan des sources de biais :

- Persistance des cadavres
- Détection des cadavres
- Biais surfacique
- « Cripping loss »





Bilan des sources de biais :

- Persistance des cadavres
- Détection des cadavres
- Biais surfacique
- « Cripping loss »

Conséquences:

- Tous conduisent à une sous-estimation des mortalités
- Leur hétérogénéité impose de les estimer localement
- Ils doivent être intégrés dans des formules pour corriger le nombre de cadavres effectivement détectés





Les quatre formules proposées par le Ministère

Formule de Winkelmann

$$N = (\frac{C}{p \times d}) \times A$$

Formule d'Erickson (2004)
$$N = (\frac{I \times C}{t \times d}) \times A$$

$$N = (\frac{I \times C}{t \times d}) \times A$$

Formule de Jones (2009)
$$N = \frac{C}{d \times exp^{-0.5 \times I/t} \times \hat{e}} \times A \text{ avec } p = exp^{(-0.5 \times \frac{I}{t})}$$

Formule de Huso (2010)
$$N = C/(d \times \frac{t \times \left(1 - exp^{-\frac{l}{t}}\right)}{I}) \times \hat{e}) \times A \ avec \ p = t \times (1 - exp^{-\frac{l}{t}})/I$$

C : Nombre de cadavres comptés ;

p: Taux de persistance. Proportion de cadavres qui perdurent au-delà de l'intervalle entre deux passages;

d : Probabilité de détection des cadavres;

A: Coefficient de correction surfacique (ici pour 7 anneaux)

$$A = \frac{\sum_{k^7} Ck / Sk}{\sum_{k^7} Ck}$$

Où S_k est la proportion prospectée du carré de 1ha et C_k est le nombre de cadavres comptés sur le carré prospecté.

: Nombre de jours entre chaque passage ;

- : Intervalle effectif = $-\log(0.01) \times t$;
- t : Nombre de jour moyen durant lequel les cadavres perdurent sur place (par défaut, les poussins si le nombre de cadavres laissés sur place n'est pas suffisant)
- ê : Coefficient correcteur de l'intervalle = Min (I : i) / I





Une hétérogénéité des résultats entre formules

- Les quatre formules conduisent à des résultats parfois très différents (parfois un facteur 10, voire même plus dans certaines situations)
- Toutes ces formules reposent sur des hypothèses de modélisation plus ou moins simplificatrices qui ne sont jamais discutées
- Plusieurs autres formules existent reposant sur d'autres hypothèses
- Certaines hypothèses sont à l'évidence trop simplificatrices (Winkelmann par exemple)...mais pour les autres...??
- Les conditions d'application de ces méthodes, ou les biais attendus dans telle ou telle situation sont globalement inconnues.





Table 1 Comparative analysis of the adjustment terms and assumptions associated with the fatality estimators regarding searched area, search interval, carcass removal and searcher detection probability.

			Potential error sources	
Estimator	Search area	Search interval	Carcass removal	Search detection probability
Erickson et al. 2000 (eq. 1/3)	Adjustment term based on the proportion of turbines searched.	No explicit requirements.	Adjustment term based on the mean persistence time (in days). Considers right-censored observations. Assumes that removal times follow an exponential distribution.	Adjustment term based on the empirical proportion of carcasses detected by the searchers. Carcass not found during the first search can be found in a subsequent search.
Shoenfeld 2004 (eq. 4/5)	Adjustment term based on the proportion of turbines searched.	In eq. 4 the number of searches is assumed to follow a Poisson model with a rate $1/\lambda$ (with λ being the mean time between searches). Eq. 5 implies regular search intervals.	Adjustment term based on the mean persistence time (in days). Considers right-censored observations. Assumes that removal times follow a Poisson distribution, with a rate $1/\bar{t}$ (eq. 4), or an exponential distribution (eq. 5).	Adjustment term based on the empirical proportion of carcasses detected by searchers. Carcass not found during the first search can be found in a subsequent search.
Kerns et al. 2005 (eq. 6/7)	Adjustment term (A, see eq. 7) accounts for the area that is not searched.	Implies regular search intervals.	Carcass persistence probability is estimated by the empirical survivor function represented by $S_T(t) = 1 - P[T \le t]$.	Detection probability estimated by distance sampling analysis. Carcass not found during the first search can be found in a subsequent search. Assumes constant carcass detectability over time.
Jain et al. 2007 (eq. 8)	Adjustment term based on the proportion of turbines searched.	No explicit requirements.	Adjustment term based on the empirical proportion of persisting carcasses after approximately half of the search interval.	Adjustment term based on the empirical proportion of carcasses detected by searchers. Carcasses overlooked are assumed to have zero probability to be detected in subsequent searches.

Table 1 (Continued)

			Potential error sources	
Estimator	Search area	Search interval	Carcass removal	Search detection probability
Pollock (2007) (eq. 8/9)	Not considered in the original formula.	Implies regular search intervals.	Adjustment term based on the empirical proportion of persisting carcasses (see eq. 10). The author claims to assume that the number of verifications until the first carcass removal occurs follows a geometric model.	Adjustment term based on the empirical proportion of carcasses detected by searchers. Carcasses overlooked are assumed to have zero probability to be detected in subsequent searches.
Huso 2010 (eq. 10)	Adjustment term based on the proportion of animals that die outside the search plot and the probability of including that plot in the sample of the turbines searched.	Considers the 'effective search interval' (ν) based on the i (the length of time beyond which the probability of a carcass persisting is $\leq 1\%$).	Adjustment term based on the mean persistence time (in days). Considers right-censored observations. Assumes that removal times follow an exponential distribution.	Adjustment term based on the empirical proportion of carcasses detected by searchers. Carcasses overlooked are assumed to have zero probability to be detected in subsequent searches.
Korner-Nievergelt et al. 2011 (eq. 11/12)	Not considered in the original formula.	Implies regular search intervals.	Adjustment term based on daily persistence probability. Carcass removal is assumed to be constant over time.	Adjustment term based on the empirical proportion of carcasses detected by searchers. Carcasses not found during the first search can be found in a subsequent search. Assumes constant (eq. 11) or decreasing (eq. 12) carcass detectability over time.

Table 1 Comparative analysis of the adjustment terms and assumptions associated with the fatality estimators regarding searched area, search interval, carcass removal and searcher detection probability.

			Potential error sources	
Estimator	Search area	Search interval	Carcass removal	Search detection probability
Erickson et al. 2000 (eq. 1/3)	Adjustment term based on the proportion of turbines searched.	No explicit requirements.	Adjustment term based on the mean persistence time (in days). Considers right-censored observations. Assumes that removal times follow an exponential distribution.	Adjustment term based on the empirical proportion of carcasses detected by the searchers. Carcass not found during the first search can be found in a subsequent search.
Shoenfeld 2004 (eq. 4/5)	Adjustment term based on the proportion of turbines searched.	In eq. 4 the number of searches is assumed to follow a Poisson model with a rate $1/\lambda$ (with λ being the mean time between searches). Eq. 5 implies regular search intervals.	Adjustment term based on the mean persistence time (in days). Considers right-censored observations. Assumes that removal times follow a Poisson distribution, with a rate $1/\bar{t}$ (eq. 4), or an exponential distribution (eq. 5).	Adjustment term based on the empirical proportion of carcasses detected by searchers. Carcass not found during the first search can be found in a subsequent search.
Kerns et al. 2005 (eq. 6/7)	Adjustment term (A, see eq. 7) accounts for the area that is not searched.	Implies regular search intervals.	Carcass persistence probability is estimated by the empirical survivor function represented by $S_T(t) = 1 - P[T \le t]$.	Detection probability estimated by distance sampling analysis. Carcass not found during the first search can be found in a subsequent search. Assumes constant carcass detectability over time.
Jain et al. 2007 (eq. 8)	Adjustment term based on the proportion of turbines searched.	No explicit requirements.	Adjustment term based on the empirical proportion of persisting carcasses after approximately half of the search interval.	Adjustment term based on the empirical proportion of carcasses detected by searchers. Carcasses overlooked are assumed to have zero probability to be detected in subsequent searches.

Table 1 (Continued)

	Potential error sources						
Estimator	Search area	Search interval	Carcass removal	Search detection probability			
Pollock (2007) (eq. 8/9)	Not considered in the original formula.	Implies regular search intervals.	Adjustment term based on the empirical proportion of persisting carcasses (see eq. 10). The author claims to assume that the number of verifications until the first carcass removal occurs follows a geometric model.	Adjustment term based on the empirical proportion of carcasses detected by searchers. Carcasses overlooked are assumed to have zero probability to be detected in subsequent searches.			
Huso 2010 (eq. 10)	Adjustment term based on the proportion of animals that die outside the search plot and the probability of including that plot in the sample of the turbines searched.	Considers the 'effective search interval' (v) based on the i (the length of time beyond which the probability of a carcass persisting is $\leq 1\%$).	Adjustment term based on the mean persistence time (in days). Considers right-censored observations. Assumes that removal times follow an exponential distribution.	Adjustment term based on the empirical proportion of carcasses detected by searchers. Carcasses overlooked are assumed to have zero probability to be detected in subsequent searches.			
Korner-Nievergelt et al. 2011 (eq. 11/12)	Not considered in the original formula.	Implies regular search intervals.	Adjustment term based on daily persistence probability. Carcass removal is assumed to be constant over time.	Adjustment term based on the empirical proportion of carcasses detected by searchers. Carcasses not found during the first search can be found in a subsequent search. Assumes constant (eq. 11) or decreasing (eq. 12) carcass detectability over time.			

Table 1 Comparative analysis of the adjustment terms and assumptions associated with the fatality estimators regarding searched area, search interval, carcass removal and searcher detection probability.

			Potential error sources	
Estimator	Search area	Search interval	Carcass removal	Search detection probability
Erickson et al. 2000 (eq. 1/3)	Adjustment term based on the proportion of turbines searched.	No explicit requirements.	Adjustment term based on the mean persistence time (in days). Considers right-censored observations Assumes that removal times follow an exponential distribution.	Adjustment term based on the empirical proportion of carcasses detected by the searchers. Carcass not found during the first search can be found in a subsequent search.
Shoenfeld 2004 (eq. 4/5)	Adjustment term based on the proportion of turbines searched.	In eq. 4 the number of searches is assumed to follow a Poisson model with a rate $1/\lambda$ (with λ being the mean time between searches). Eq. 5 implies regular search intervals.	Adjustment term based on the mean persistence time (in days). Considers right-censored observations. Assumes that removal times follow a Poisson distribution, with a rate $1/\bar{t}$ (eq. 4), or an exponential distribution (eq. 5).	Adjustment term based on the empirical proportion of carcasses detected by searchers. Carcass not found during the first search can be found in a subsequent search.
Kerns et al. 2005 (eq. 6/7)	Adjustment term (A, see eq. 7) accounts for the area that is not searched.	Implies regular search intervals.	Carcass persistence probability is estimated by the empirical survivor function represented by $S_T(t) = 1 - P[T \le t]$.	Detection probability estimated by distance sampling analysis. Carcass not found during the first search can be found in a subsequent search. Assumes constant carcass detectability over time.
Jain et al. 2007 (eq. 8)	Adjustment term based on the proportion of turbines searched.	No explicit requirements.	Adjustment term based on the empirical proportion of persisting carcasses after approximately half of the search interval.	Adjustment term based on the empirical proportion of carcasses detected by searchers. Carcasses overlooked are assumed to have zero probability to be detected in subsequent searches.

L'estimation des mortalités, un processus statistique

- Les facteurs de correction sont des estimateurs
- Ils sont associés à un certain degré d'erreur car issus d'un processus aléatoire





L'estimation des mortalités, un processus statistique

- Les facteurs de correction sont des estimateurs
- Ils sont associés à un certain degré d'erreur car issus d'un processus aléatoire

Exemple de la vie courante :

- Le sexe-ratio chez l'homme est de 0.5
- Si on prend une famille au hasard à Montpellier on peut très bien tomber du 4 garçons et 0 filles le sexe-ratio est alors estimé à 1!
- On peut calculer la probabilité d'avoir 4 garçons dans une fratrie de 4 enfants, elle est de 0.06, ça n'est donc pas très fréquent, mais pas totalement improbable.



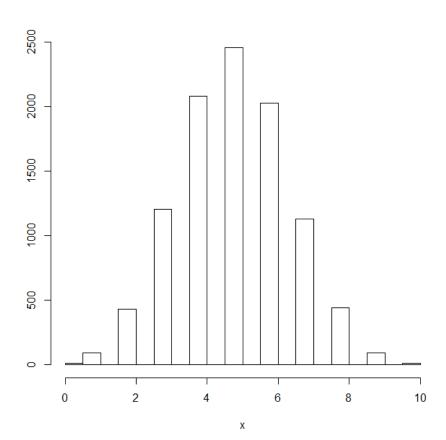


10.000 tirages aléatoires parmi 10 cadavres avec une probabilité de 0,50 de trouver chaque cadavre





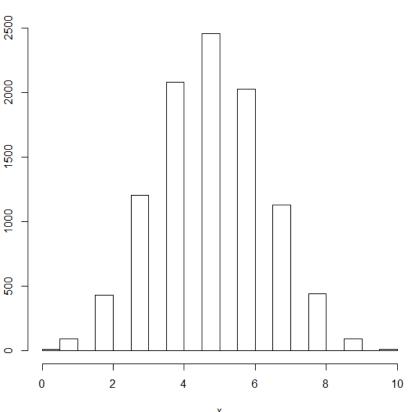
10.000 tirages aléatoires parmi 10 cadavres avec une probabilité de 0,50 de trouver chaque cadavre







10.000 tirages aléatoires parmi 10 cadavres avec une probabilité de 0,50 de trouver chaque cadavre



Moyenne estimée: 0,50 (ouf)

Intervalle de confiance à 95% : [0,20-0,80]

Valable pour la détection ET la persistance.



Les deux sources d'incertitude se « cumulent ».



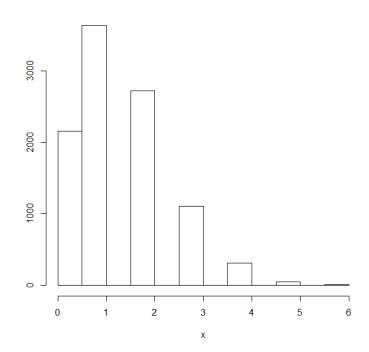


10.000 tirages aléatoires avec 50 cadavres répartis sur toute l'année, la probabilité quotidienne d'avoir une mortalité : 0,14
Si on échantillonne 10 jours...





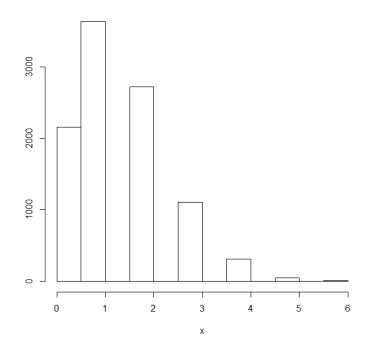
10.000 tirages aléatoires avec 50 cadavres répartis sur toute l'année, la probabilité quotidienne d'avoir une mortalité : 0,14
Si on échantillonne 10 jours...







10.000 tirages aléatoires avec 50 cadavres répartis sur toute l'année, la probabilité quotidienne d'avoir une mortalité : 0,14
Si on échantillonne 10 jours...



Dans 20% des cas on ne trouve aucun cadavre. Le nombre varie de 0 à 6.

Avec 10 mortalités sur l'année, dans 76% des cas il n'y a pas de cadavre.

Seule une (très) forte persistance assure alors de détecter quelques cadavres (cumul de plusieurs jours).





Des « erreurs standards » (ou intervalles de confiance) à propager

- > Trois paramètres estimés :
 - La probabilité de détection
 - La persistance
 - Le nombre moyen de cadavres trouvés par jour de prospection

Dont la précision dépend de la taille de l'échantillon et de l'hétérogénéité du phénomène.





Des « erreurs standards » (ou intervalles de confiance) à propager

- > Trois paramètres estimés :
 - La probabilité de détection
 - La persistance
 - Le nombre moyen de cadavres trouvés par jour de prospection

Dont la précision dépend de la taille de l'échantillon et de l'hétérogénéité du phénomène.

➤ Il faut combiner toutes les (im)précisions de ces estimations pour obtenir la précision globale à travers les formules proposées.





Des « erreurs standards » (ou intervalles de confiance) à propager

- > Trois paramètres estimés :
 - La probabilité de détection
 - La persistance
 - Le nombre moyen de cadavres trouvés par jour de prospection

Dont la précision dépend de la taille de l'échantillon et de l'hétérogénéité du phénomène.

➤ Il faut combiner toutes les (im)précisions de ces estimations pour obtenir la précision globale à travers les formules proposées.

Mais:

- pas de calcul analytique
- La calcul des intervalles de confiance des différentes formules implique du « boostraping » (on répète 10.000 fois le calcul de la formule en prenant des valeurs aux hasard dans la distribution de chaque estimation)





Nb de passages	100	100	50	50	12	12
Intervalle entre passages	4	4	4	4	4	4
Persistance quotidienne	0,8	0,5	0,8	0,5	0,8	0,5
Probabilité de détection	0,5	0,5	0,5	0,5	0,5	0,5
Nb de mortalités annuelle	36	36	36	36	36	36
Estimation de la mortalité	59	58	58	61	61	60
intervalle confiance à 95% inf	28	14	22	7	0	0
intervalle confiance à 95% sup	111	137	122	172	160	244





Nb de passages	100	100	50	50	12	12
Intervalle entre passages	4	4	4	4	4	4
Persistance quotidienne	0,8	0,5	0,8	0,5	0,8	0,5
Probabilité de détection	0,5	0,5	0,5	0,5	0,5	0,5
Nb de mortalités annuelle	36	36	36	36	36	36
Estimation de la mortalité	59	58	58	61	61	60
intervalle confiance à 95% inf	28	14	22	7	0	0
intervalle confiance à 95% sup	111	137	122	172	160	244





Nb de passages	100	100	50	50	12	12
Intervalle entre passages	4	4	4	4	4	4
Persistance quotidienne	0,8	(0,5)	0,8	0,5	0,8	0,5
Probabilité de détection	0,5	0,5	0,5	0,5	0,5	0,5
Nb de mortalités annuelle	36	36	36	36	36	36
Estimation de la mortalité	59	58	58	61	61	60
intervalle confiance à 95% inf	28	14	22	7	0	0
intervalle confiance à 95% sup	111	137	122	172	160	244





Nb de passages	100	100	50	50	12	12
Intervalle entre passages	4	4	4	4	4	4
Persistance quotidienne	0,8	0,5	0,8	0,5	0,8	0,5
Probabilité de détection	0,5	0,5	0,5	0,5	0,5	0,5
Nb de mortalités annuelle	36	36	36	36	36	36
Estimation de la mortalité	59	58	58	61	61	60
intervalle confiance à 95% inf	28	14	22	7	0	0
intervalle confiance à 95% sup	111	137	122	172	160	244





Nb de passages	100	100	50	50	12	12
Intervalle entre passages	4	4	4	4	4	4
Persistance quotidienne	0,8	0,5	0,8	0,5	0,8	0,5
Probabilité de détection	0,5	0,5	0,5	0,5	0,5	0,5
Nb de mortalités annuelle	36	36	36	36	36	36
Estimation de la mortalité	59	58	58	61	61	60
intervalle confiance à 95% inf	28	14	22	7	0	0
intervalle confiance à 95% sup	111	137	122	172	160	244





Si on cumule toutes les sources d'incertitudes...(ici avec Huso 2010)

Nb de passages	100	100	50	50	12	12
Intervalle entre passages	4	4	4	4	4	4
Persistance quotidienne	0,8	0,5	0,8	0,5	0,8	0,5
Probabilité de détection	0,5	0,5	0,5	0,5	0,5	0,5
Nb de mortalités annuelle	36	36	36	36	36	36
Estimation de la mortalité	59	58	58	61	61	60
intervalle confiance à 95% inf	28	14	22	7	0	0
intervalle confiance à 95% sup	111	137	122	172	160	244

Il faut que la persistance soit forte ou que les mortalités soient très nombreuses.

Par exemple : 1 cadavres par jour et persistance de 0,8 avec 100 visites on peut espérer un facteur 2 sur les bornes de l'intervalle de confiance.



