

Alose_ABC : une application pour estimer le nombre d'aloses femelles à partir du dénombrement de leurs bulls

MODE D'EMPLOI

Table des matières

Introduction.....	2
Modes d'exécution de l'application	2
En local, sur R	2
Exécution en local sur R du script en ligne sur GitHub.....	3
Exécution en ligne	3
Utilisation	4
Onglet "Données"	4
Onglet "Paramètres"	6
Onglet "Résultats"	8

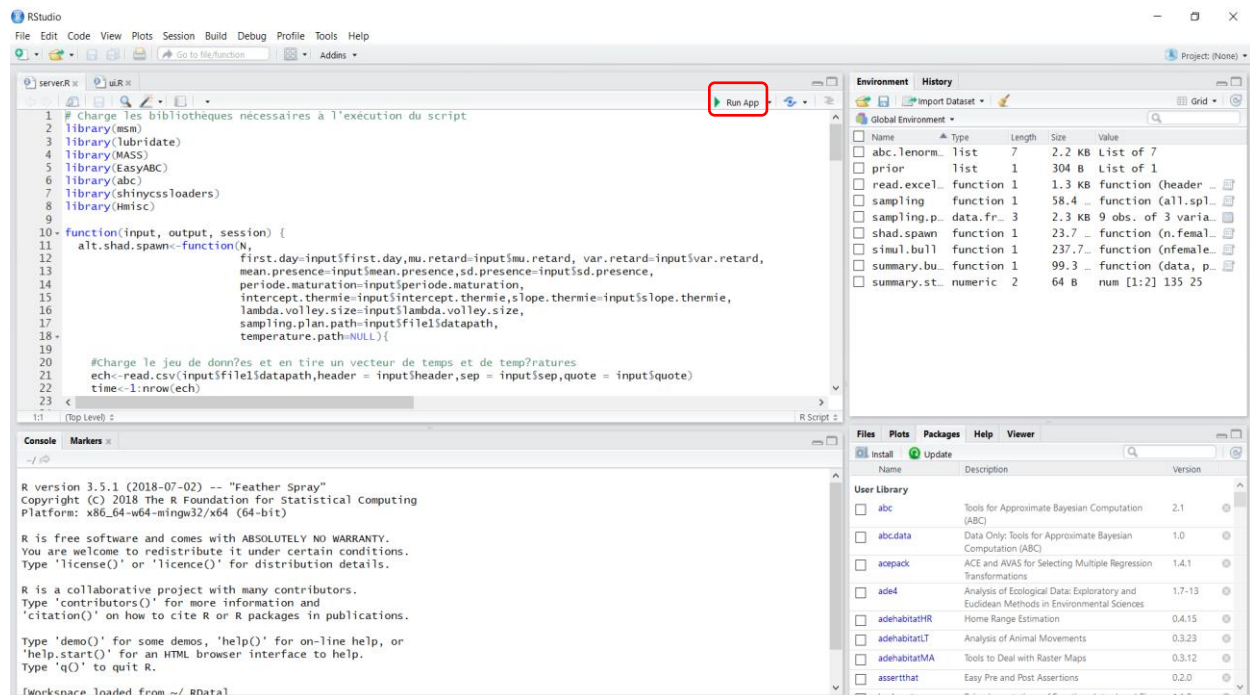
Introduction

L'application "Alose_ABC" permet d'estimer le nombre d'aloses femelles sur une frayère à partir du nombre de bulls détectés lors de chacune des nuits échantillonnées au cours de la saison de reproduction, de la température de l'eau lors de chaque nuit de la saison, et des bornes inférieure et supérieure *a priori* de l'effectif. En se basant sur des hypothèses concernant le comportement reproducteur des aloses (les paramètres du modèle, dont les valeurs par défaut ont été inférées par observation du comportement individuel dans la Nivelle en 2017), l'application utilise un algorithme ABC (Approximate Bayesian Computation) pour inférer une distribution *a posteriori* du nombre de femelles. Ce document est le mode d'emploi de l'application. Il ne détaille pas le contenu du modèle, dont le code source est disponible par ailleurs (https://github.com/CedricTentelier/Alose_ABC).

Modes d'exécution de l'application

En local, sur R

Le code source de l'application (fichiers server.R et ui.R) peut être téléchargé sur https://github.com/CedricTentelier/Alose_ABC. L'application peut alors être utilisée en local, sur un poste où sont installés les logiciels R (<https://cran.r-project.org/>) et Rstudio (<https://www.rstudio.com/>), ainsi que les bibliothèques R nécessaires à l'exécution de l'application, dont la bibliothèque Shiny (<https://www.rstudio.com/products/shiny/>) qui permet de développer et exécuter une application à partir d'un script R. Les fichiers server.R et ui.R doivent être enregistrés sur l'ordinateur dans un dossier commun, pour être ensuite ouverts avec Rstudio, d'où l'application peut être lancée en cliquant sur le bouton "Run App" en haut à droite de la fenêtre du script.



Exécution en local sur R du script en ligne sur GitHub

Depuis Rstudio, la commande

```
library(shiny)
runGitHub(repo="Alose_ABC", username="CedricTentelier")
```

permet d'accéder en ligne au script présent sur GitHub et de l'exécuter en local sur R. L'avantage de cette méthode par rapport à la précédente est qu'en cas de mise à jour, l'utilisateur accède systématiquement à la dernière version du script sans avoir à la télécharger.

Exécution en ligne

L'utilisateur peut exécuter l'application complètement en ligne, en entrant l'adresse suivante dans son navigateur : https://ctentelier.shinyapps.io/alt_shad_ABC/

L'avantage de cette méthode est non seulement d'accéder systématiquement à la dernière mise à jour de l'application mais aussi de pouvoir l'exécuter sans disposer de R ou Rstudio sur son poste de travail. En effet l'application est exécutée sur un serveur distant, avec lequel l'interface web permet d'échanger les informations nécessaires (téléchargement des données, réglage des paramètres, visualisation des résultats).

Utilisation

Quel que soit le mode d'exécution, l'application *Alose_ABC* s'ouvre dans le navigateur web de l'utilisateur, et comporte trois onglets : "Données", "Paramètres" et "Résultats".

Onglet "Données"

Cet onglet contient deux encadrés, permettant à l'utilisateur de transmettre à l'application les données observées.

Bulls observés et température de l'eau

Fichier csv dont la colonne 1 contient la date des nuits de toute la saison de reproduction (JJ/MM/AAAA), la colonne 2 contient le nombre de bulls détectés à chaque nuit échantillonnée (NA si la nuit n'a pas été échantillonnée), la colonne 3 contient la température de l'eau pour chaque nuit.

Choisissez le fichier csv

Browse... audio_&_temperature_2018.csv

Upload complete

☒ Mes données ont des en-têtes

Séparateur

☐ Virgule

☒ Point virgule

☐ Espace

Marqueur de décimale

☐ Virgule

☒ Point

Guillemets

☐ Pas de guillemets

☒ Guillemets doubles

☐ Guillemets simples

Date du premier jour de reproduction de la population (aaaa-mm-jj)

2018-04-25

Nombre minimal d'aloses, a priori

1

Nombre maximal d'aloses, a priori

200

date	bulls	temperature
24/04/2018	0	16.40
25/04/2018	1	14.30
26/04/2018	0	13.70

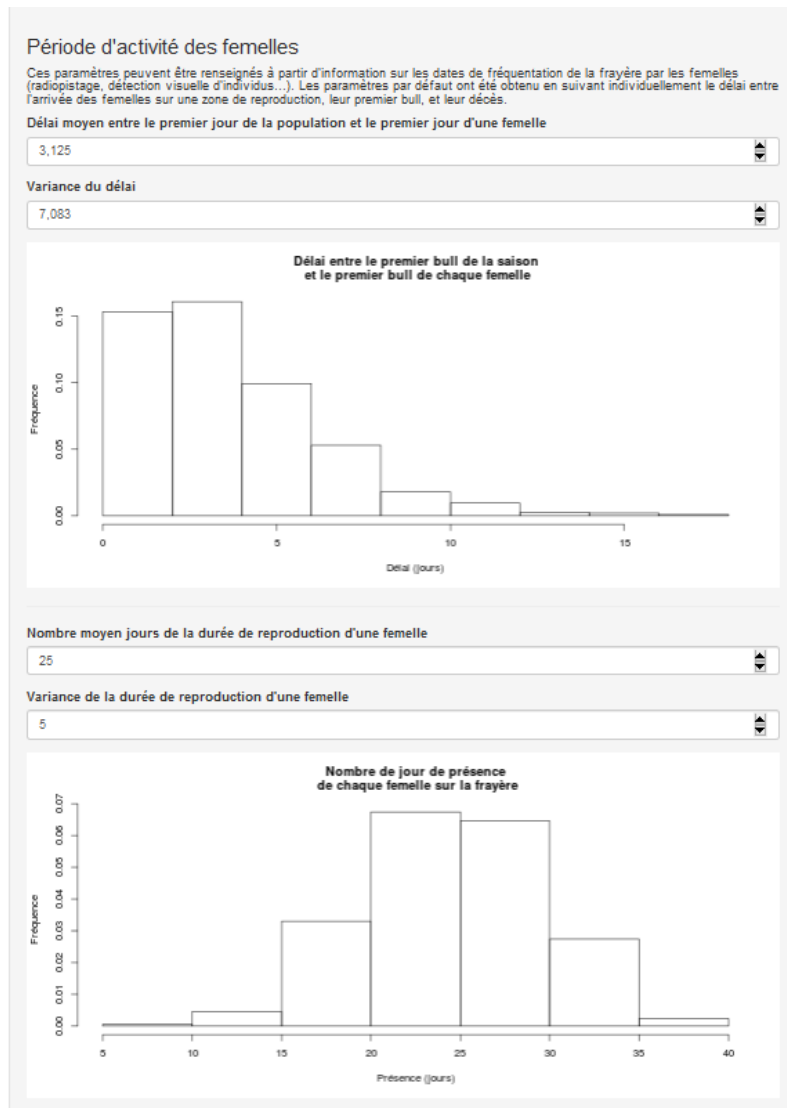
Dans l'encadré de gauche, l'utilisateur télécharge son jeu de données. Il s'agit d'un tableau au format .csv comportant trois colonnes et dont la mise en forme (en-têtes, séparateur de colonnes, marqueur de décimales, guillemets) peut être spécifiée. La première colonne correspond à la date de chaque nuit de la saison de reproduction. La date est entrée au format JJ/MM/AAAA (jour, mois, année). La date est celle du soir où commence la nuit. Ainsi les données collectées dans la nuit du 23/05 au 24/05 sont reportées à la date du 23/05. La deuxième colonne correspond au nombre de bulls détectés au cours de chaque nuit. On suppose ici qu'une nuit est soit complètement échantillonnée (de 22h00 à 6h00) soit pas du tout. **Si une nuit n'a pas été échantillonnée, il est important de noter "NA" (pour *not available*) et pas zéro**, pour que le modèle distingue une nuit non échantillonnée d'une nuit échantillonnée mais lors de laquelle aucun bull n'a été détecté. La troisième colonne correspond à la température de l'eau au cours de chaque nuit. On suppose que cette température est disponible pour toutes les nuits de la saison, même si les bulls n'ont pas été échantillonnés à chaque nuit. Une fois téléchargé, le jeu de données apparaît sous l'encadré.

Dans l'encadré de droite, l'utilisateur saisit des informations complémentaires, nécessaire à l'exécution du modèle. La première de ces informations est la date du premier jour de reproduction sur la frayère, c'est-à-dire du premier bull. Si l'écoute des bulls a commencé suffisamment tôt, cette date peut être la première du jeu de données téléchargé pour laquelle le nombre de bulls est non nul. Cependant, il se peut que l'utilisateur dispose d'informations complémentaires lui indiquant qu'un bull s'est produit à une certaine date, avant le début de l'échantillonnage formel. Les deux autres informations à

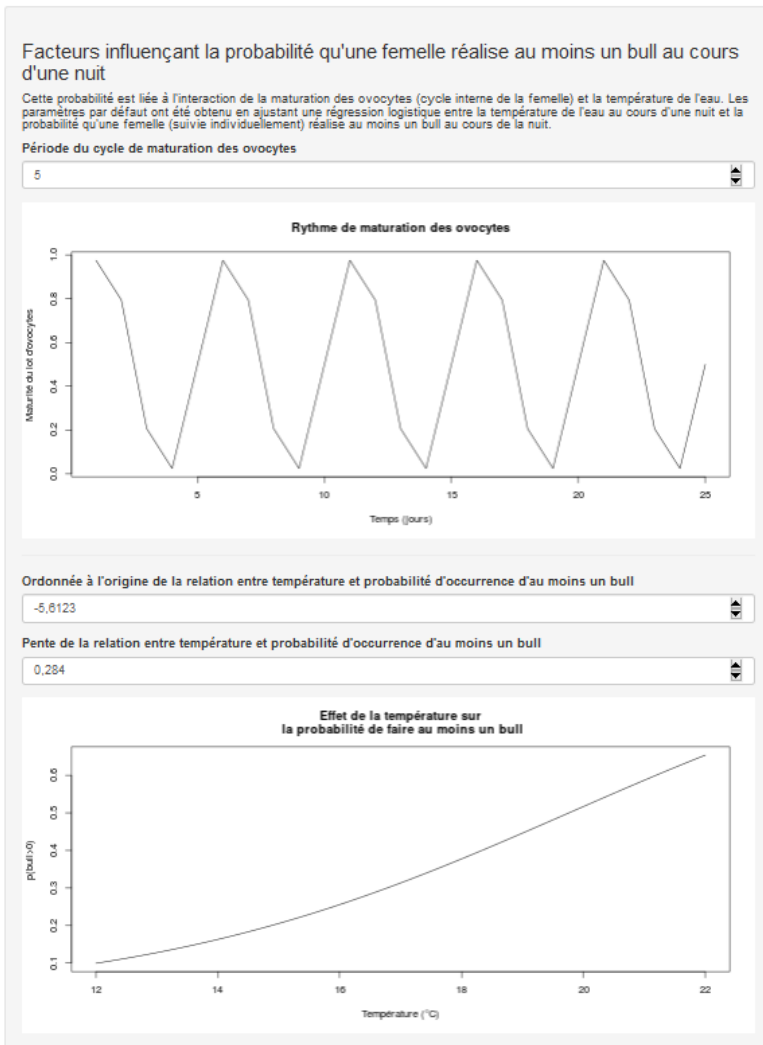
renseigner sont les nombres *a priori* minimum et maximum de femelles. Ces valeurs *a priori* doivent être indépendantes du jeu de données téléchargé. Elles peuvent par exemple être issues de comptages partiels sur des obstacles à l'aval de la frayère, ou du minimum et du maximum observés lors des années précédentes. Elles contraignent le modèle dans son inférence de l'effectif. Ainsi des bornes trop étroites restreignent le modèle, donc limitent l'utilisation des données ; des bornes trop larges associées à des données trop partielles laissent la possibilité à une grande gamme de valeurs même *a priori* improbables.

Onglet "Paramètres"

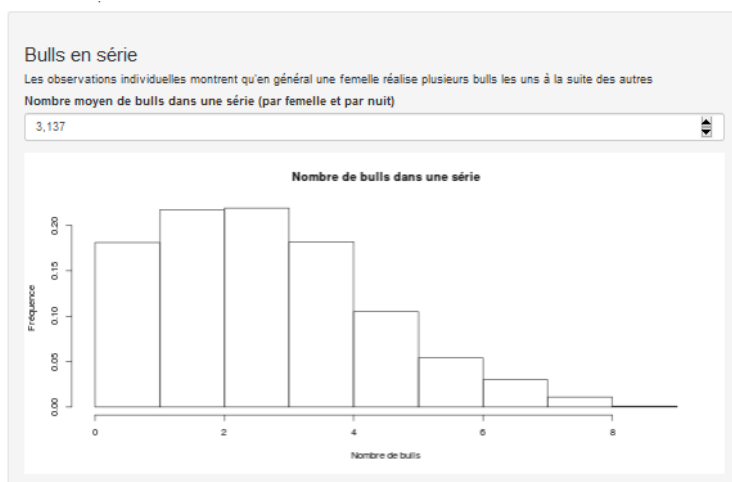
Cet onglet contient trois encadrés permettant à l'utilisateur, de consulter et éventuellement modifier les paramètres du modèle d'estimation. Ces paramètres concernent le comportement reproducteur individuel des aloses femelles, est ont été estimé à partir d'observations individuelles (avec radio pistage et accélérométrie) menées sur huit femelles dans la Nivelle en 2017. Etant donné que ces paramètres individuels sont très rarement accessibles, il est conseillé de garder les valeurs par défaut, à moins de disposer d'informations justifiant de les modifier.



Le premier encadré concerne la période d'activité des femelles sur la frayère. L'utilisateur peut consulter et modifier la moyenne et la variance du délai (en jours) entre le début de la saison de reproduction (premier bull, toutes femelles confondues) et l'entrée en activité de chaque femelle. Ce délai est tiré dans une loi Binomiale Négative, dont la distribution est représentée dans l'encadré et se met à jour si les paramètres sont modifiés. De la même manière, l'utilisateur peut consulter et modifier la moyenne et la variance de la durée (en jours) pendant laquelle une femelle est active (entre son premier bull et son décès). Cette durée suit une distribution Normale, dont la distribution est représentée dans l'encadré et se met à jour si les paramètres sont modifiés.



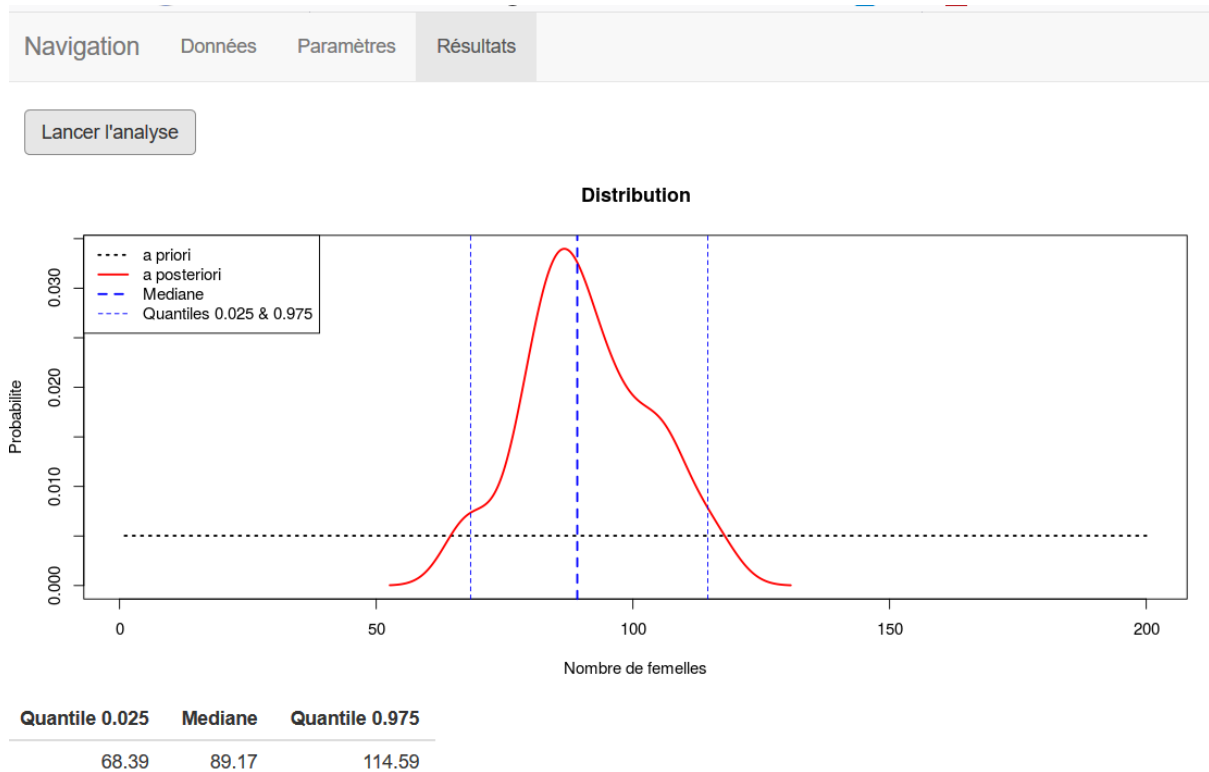
Le deuxième encadré de l'onglet "Paramètres" concerne les facteurs influençant la probabilité qu'une femelle réalise au moins un bull au cours d'une nuit. Cette probabilité est le produit de deux fonctions représentant, pour l'une le rythme de maturation des ovocytes de la femelle, et pour l'autre l'effet de la température de l'eau sur le déclenchement des bulls. La fonction de maturation est une sinusoïde dont la période est ajustable. La période par défaut de 5 jours correspond à la maturation séquentielle des ovocytes, dont chaque lot semble être expulsé tous les 5 jours en moyenne. La fonction liant la température à la probabilité d'occurrence d'un bull représente une régression logistique dont l'ordonnée à l'origine et la pente (effet de la température) peuvent être modifiées. Les deux fonctions sont représentées par des graphiques qui se remettent à jour quand les paramètres sont modifiés.



Le troisième encadré de l'onglet "Paramètres" correspond au nombre de bulls qu'une même femelle réalise au cours d'une même nuit. Les observations individuelles ont montré qu'une femelle réalisait en général plusieurs bulls au cours d'une même nuit, et que le nombre de bull dans la série suivait une loi de Poisson de paramètre (moyenne et variance) 3.137. Si ce paramètre est modifié, l'histogramme représentant la distribution est mis à jour.

Onglet "Résultats"

Une fois les données entrées et les valeurs de paramètres vérifiées ou modifiées, l'utilisateur se rend sur l'onglet "Résultats" pour lancer l'analyse en cliquant sur le bouton "Lancer l'analyse", et visualiser les résultats affichés une fois que l'analyse est terminée, ce qui peut prendre quelques secondes.



Le graphique qui s'affiche à la fin de l'analyse représente la distribution *a priori* de l'effectif de femelles (une loi uniforme bornée entre le minimum et le maximum défini par l'utilisateur), ainsi que la distribution *a posteriori* inférée par le modèle. Sur cette distribution *a posteriori* sont repérés la médiane et les quantiles à 2.5% et 97.5%, qui peuvent être considérés comme les bornes de l'intervalle de crédibilité à 95% de l'estimation. Plus ces bornes sont étroites, plus l'incertitude de l'estimation est réduite. La médiane et les quantiles à 2.5% et 97.5% sont également représentés dans un tableau sous le graphique.

Pour lancer une nouvelle analyse, il suffit de cliquer à nouveau sur le bouton correspondant, éventuellement après avoir modifié les données ou les paramètres sur les autres onglets.