

Ce tutoriel permet de réaliser de manière automatisée des analyses régionales de variations temporelles d'abondances de populations d'oiseaux communs suivies par le programme STOC-EPS. Contrairement au logiciel TRIM, l'analyse via R permet d'enregistrer directement les tendances et graphiques correspondant pour chaque espèce, ainsi que les indicateurs par groupes d'espèces spécialistes par habitats.

Informations préliminaires

Le dossier que vous avez téléchargé (ici : <http://vigienature.mnhn.fr/page/protocole-et-aide-lanalyse>) contient plusieurs éléments :

- « **R-3.2.2-win.exe** » : Il s'agit du logiciel R, qui va vous permettre d'effectuer vos analyses de tendances. Il faudra l'installer sur votre ordinateur en double cliquant sur le fichier, puis en suivant les instructions.
- « **Commencer.RData** » : C'est le fichier que vous devrez ouvrir pour pouvoir débiter les analyses.
- « **scirptSTOCeps.r** » : C'est le script qu'utilise le logiciel R pour faire les analyses de vos tendances. Il ne faut pas modifier ou déplacer ce fichier qui contient toutes les lignes de commande en langage R.
- « **TutorielAnalyses.pdf** » : Il s'agit de ce tutoriel d'analyse des tendances du STOC-EPS, qui détaille toutes les étapes nécessaires à la construction et l'analyse de vos tendances.
- « **Donnees** » : C'est le dossier où vous devrez mettre vos données STOC-EPS à analyser. Les fichiers de données peuvent être en format « .txt » ou « .csv », et ne doivent pas comporter d'espace ou de caractère spéciaux (de type : &,é,ï,-,...)
- « **Librairie** » : Ce dossier contient des éléments de références (tels que les codes espèces) qui vont servir à analyser vos fichiers de données.
- « **Resultats** » : C'est dans ce dossier que seront stockés les résultats des analyses de tendances, en format « .csv » ouvrables avec Excel, et en format « .png » pour les graphiques, ouvrables avec la visionneuse d'image Windows où n'importe quel autre logiciel de traitement d'image.

Etapes à réaliser avant de lancer les analyses

1) Préparation du fichier de données :

Avant de lancer les analyses de tendances, il vous faut créer votre fichier de données STOC-EPS. Pour rappel ce fichier doit contenir par carré le maximum d'individu observé sur les 2 passages (maximum sommé pour les 10 points du

carré). Ce fichier peut être de la forme que vous le souhaitez, mais devra respecter quelques contraintes afin de pouvoir être analysable sous R :

- Être en format « .txt » ou « .csv »
- Ne comporter qu'une seule feuille
- Ne pas comporter d'espaces, d'accents (é,è,à,ê,ï,ù) ou de caractères spéciaux («&,@,-,;,/»). Afin d'être sûr d'éliminer tous ces éléments, sous Excel, sélectionnez la totalité de la feuille, puis allez dans l'onglet « Remplacer », puis dans « Rechercher : » tapez juste un espace (pour rechercher tous les espaces dans votre document), et dans « Remplacer par : » ne mettez rien, puis « OK ». Procédez de même pour les accents et caractères spéciaux susceptibles d'être dans votre document.
- Avoir la première ligne comme ligne d'en-tête, avec le nombre de colonne que vous souhaitez, mais au moins les colonnes « annee » et « carre », puis au choix selon la façon dont vous organisez vos données, une colonne par espèces avec les codes espèces du STOC (voir la liste de référence dans le dossier « Librairie »), c'est-à-dire sous cette forme :

	A	B	C	D	E	F	G	H	I	J	K	L
1	carre	dept	region	annee	x	y	ALAARV	ALERUF	ANTCAM	ANTPRA	APUAPU	BUTBU
2	CARRE010001	1	rhonealpes	2014	435987	9987346	0	0	0	0	0	
3	CARRE010002	1	rhonealpes	2013	435998	9954677	0	0	0	0	0	2
4	CARRE010003	1	rhonealpes	2012	805072	2134026	0	0	0	0	0	0
5	CARRE010004	1	rhonealpes	2011	876543	9876543	0	2	0	0	0	4
6	CARRE010001	1	rhonealpes	2014	435987	9987346	0	0	0	0	0	0
7	CARRE010002	1	rhonealpes	2013	435998	9954677	0	0	0	0	0	5

ou bien avec une colonnes « espece » et une colonne « nombre » sous cette forme :

	A	B	C	D	E	F	G	H
1	carre	dept	region	annee	x	y	espece	nombre
2	CARRE010001	1	rhonealpes	2014	435987	9987346	ALAARV	0
3	CARRE010001	1	rhonealpes	2014	435987	9987346	CARCHL	3
4	CARRE010001	1	rhonealpes	2014	435987	9987346	BUTBUT	1
5	CARRE010004	1	rhonealpes	2011	876543	9876543	ANTCAM	0
6	CARRE010004	1	rhonealpes	2011	876543	9876543	BUTBUT	0
7	CARRE010004	1	rhonealpes	2011	876543	9876543	CARCAN	2

Une fois ce fichier créé, vous devez le placer dans le dossier «**Donnes** » qui se trouve dans le package que vous avez téléchargé. Vous pouvez mettre autant de fichiers que vous souhaitez dans ce dossier, mais l'analyse réalisée sous R se fera par défaut sur la totalité des données des fichiers présents dans le dossier (agrégation automatique des fichiers sous R, mais possibilité de changer ce paramètre via les options de codes décrites plus loin dans le tutoriel).

2) Installation du Logiciel R :

Les analyses de tendances seront réalisées via le logiciel de statistique R. Pour l'installer, double-cliquez sur le fichier « **R-3.2.2-win.exe** », ou bien téléchargez-le à l'adresse suivante : <http://cran.univ-lyon1.fr/>.

Vous pouvez maintenant passer à l'étape suivante !

Lancement du script R

Pour débiter les analyses de vos fichiers de données, double-cliquez sur le fichier « **Commencer.RData** ».

Une fenêtre « Rconsole » s'est ouverte, c'est là que vous allez lancer l'analyse de données de vos fichiers. Les lignes de commandes que vous tapez s'afficheront en rouge, et les commentaires et messages d'informations, d'avertissements ou d'erreurs en bleu.

Afin de charger le script d'analyse de données, tapez la ligne de code suivante, puis tapez « entrée » :

```
source("scriptSTOCeps.r ")
```

Des messages d'information peuvent apparaître en bleu, vous informant par exemple du téléchargement de certains packages R.

La première fois que vous lancez le script, vous devez avoir une connexion internet afin que R puisse télécharger des « packages » qui sont utilisés dans le script.

Ensuite vous pouvez lancer les analyses en tapant la ligne de commande, puis « entrée » :
analyse()

L'analyse débute, vous pouvez suivre en direct où en est la procédure (*1) IMPORTATION*, *2) ANALYSE* ou *3) DYNAMIQUE PAR GROUPE D'ESPECES*), ainsi que les messages d'informations :

```
> analyse()

20-10-2015 14H56

1) IMPORTATION
-----

1 fichier(s)
<-- Donnees/dataTest.csv

ATTENTION COLONNE NON UTILISEE LORS DE L IMPORTATION !!!
verifiez qu il n y a pas de code espece mal orthographie

X

118 especes conservees pour l'analyse

  Code_espece          nom_espece
1     ACCNIS          Epervier d'Europe
2     ACRSCI          Rousserolle effarvate
3     AEGCAU          Mésange à longue queue
4     ALAARV          Alouette des champs
5     ALCATT          Martin-pêcheur d'Europe
```

Lors de l'étape 1) *IMPORTATION*, il est détaillé les espèces non prises en compte dans l'analyse (car distribution au cours du temps trop faible ou code espèce non reconnu), ainsi que les espèces retenues.

Des messages d'avertissements peuvent apparaître à la fin de l'analyse de type: "There were 50 or more warnings (use warnings() to see the first 50)". Ne vous inquiétez pas, c'est normal ! Veuillez ne pas en tenir compte.

Des messages d'erreurs dans l'étape « 1) *IMPORTATION* » peuvent apparaître si R n'arrive pas à lire votre fichier de données :

```
> analyse()

20-10-2015 15H16

1) IMPORTATION
-----

1 fichier(s)
<-- Donnees/dataTest.csv
Error in `[.data.frame`(data, , idvar) : undefined columns selected
> |
```

Si ce cas se présente, il est très probable qu'un espace ou un caractère spécial soit encore dans votre fichier de données ! Rouvrez-le dans Excel afin de trouver où l'erreur s'est glissée à l'aide de la fonction « Rechercher/Remplacer par ».

Il est également possible de modifier certains paramètres de l'analyse par défaut. Au lieu de taper simplement la commande `analyse()`, voici les éléments que vous pouvez paramétrer (en bleu dans le code ci-dessous, à changer selon vos souhaits) :

`analyse(listSp=c("ALAARV","BUTBUT"))` lance l'analyse uniquement sur la ou les espèces dont vous spécifiez ici le code espèce. Attention, dans ce cas l'analyse par groupe d'espèces spécialistes se fera uniquement sur les espèces indicatrices sélectionnées !

`analyse(id_session="Rhône")` donne un identifiant à l'analyse, qui va enregistrer tous les résultats dans un dossier « Rhône ».

`analyse(annees=c(2005,2013))` lance l'analyse sur la fenêtre temporelle de 2005 à 2013

`analyse(estimateAnnuel=FALSE, id_session="Rhône")` lance l'analyse uniquement des indicateurs par groupes d'espèces spécialistes et généralistes. Nécessite de rappeler de quel identifiant session il s'agit pour aller chercher les indicateurs par espèce préalablement calculés.

`analyse(figure=FALSE)` l'analyse ne fait pas les représentations graphiques des indicateurs par espèce, mais uniquement les fichiers textes.

`analyse(sauvegardeDonnees=TRUE)` dans le cas où plusieurs fichiers sont présents dans le dossier « Donnees », sauvegarde le fichier de données totale sur lequel l'analyse est faite.

`analyse(concatFile=FALSE)` lance une analyse par fichier présent dans le dossier « Données », ne concatène pas en un seul jeu de données.

analyse(description=FALSE) dans la représentation graphique de chaque espèce, seules les variations d'abondances sont représentées.

analyse(tendanceSurFigure=FALSE) les tendances, intervalles de confiances et pourcentages de variation ne sont pas affichés sur les graphiques.

analyse(tendanceGroupSpe=FALSE) le calcul des indicateurs par groupes de spécialisation par habitat n'est pas réalisé.

analyse(ICfigureGroupeSp= FALSE) les intervalles de confiance ne sont pas affichés sur le graphique de l'indicateur par groupe de spécialistes.

Vous pouvez également combiner ces différents paramètres, par exemple en tapant :

```
analyse (listSp=c("PARCAE", "PARMAJ", "PARCRI", "PARATE", "PARPAL" ),  
annees=c(2002,2012),id_session="10ans")
```

Cette ligne de code vous fera l'analyse de 5 espèces de mésange, de l'année 2002 à 2012 et sera enregistré dans un dossier nommé « 10ans ».

Astuce

Avant de lancer vos analyses sur la totalité de vos données, nous vous conseillons de lancer une pré-analyse, afin de voir la distribution de votre jeu de données.

En effet, il est possible que l'effort d'échantillonnage soit très disparate selon l'année, les années de lancement du STOC-EPS dans votre région ayant peut être beaucoup moins de carrés prospectés que les années suivantes. Afin de voir la distribution du nombre de carrés effectué dans le temps, lancez l'analyse comme expliqué précédemment, mais après que la première espèce ait été analysée, faites « échap ».

L'analyse s'arrête. Allez ensuite voir le graphique qui a été créé dans le dossier « Résultats ». La courbe verte foncée vous montre le nombre de carrés prospectés en fonction de l'année. Selon cette répartition, vous pouvez choisir de lancer l'analyse sur toutes vos données, en tapant normalement :

```
analyse()
```

ou bien sur une sélection d'années si vous constatez une grande rupture dans le nombre de carrés prospectés au cours du temps (par exemple ici pour une analyse de 2006 à 2014) :

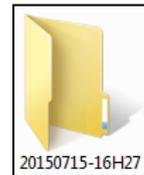
```
analyse(annees=c(2006,2014))
```

Fichiers créés dans le dossier « Analyses »

A chaque analyse que vous lancez dans R (c'est-à-dire à chaque fois que vous tapez la commande **analyse()**), un dossier se crée dans le dossier « Resultats ». Il a pour nom par

défaut la date à laquelle vous avez lancé l'analyse (format année-mois-jour), suivi de l'heure :

A l'intérieur de ce dossier sont enregistrés les résultats des analyses, à savoir :



- Un fichier image pour chaque espèce, enregistré sous la forme « **code espèce_date-heure de l'analyse .png** ». Ce fichier se compose de 3 graphiques par espèce sur la période de temps considérée : les variations d'abondance, les occurrences de l'espèce et des carrés prospectés, et enfin les abondances brutes.
- Un fichier « **tendanceGlobalEspèce.csv** », qui comporte pour chaque espèce sur la période de temps considérée les sorties du modèle statistique, le pourcentage de variation d'abondance des effectifs, la catégorie de tendance dans laquelle l'espèce se trouve (forte augmentation/augmentation modérée/stable/incertain/déclin modéré/fort déclin) et le niveau de robustesse de l'indicateur.
- Un fichier « **variationsAnnuellesEspèce.csv** », qui comporte pour chaque espèce et chaque année les sorties du modèle statistique.
- Un dossier « **Incertain** » dans lequel se trouvent les graphiques des espèces dont l'analyse n'est pas considérée comme assez robuste pour utiliser l'indicateur sans risque d'erreur trop élevé.
- Un fichier « **tendancesGlobalesGroupes.csv** », qui comporte les valeurs de variations d'abondance par groupe d'espèces de spécialistes par habitat sur la période de temps considérée ainsi que le nombre d'espèces utilisé pour le calcul de l'indicateur.
- Un fichier « **variationsAnnuellesGroupes.csv** » qui comporte pour chaque groupe de spécialistes et chaque année les valeurs de l'indicateur.
- Un fichier image « **variationsAnnuellesGroupes_date.png** », qui est la représentation graphique des variations d'abondances par groupes de spécialistes.
- Un fichier « **donneesGroupes.csv** », qui comporte pour chaque espèce et chaque année les informations relatives à leurs robustesses et donc aux poids qu'elles vont prendre dans l'analyse par groupe de spécialistes.

Lorsque vous ouvrez pour la première fois les fichiers de résultats « .csv », il est possible que selon votre configuration d'Excel, ceux-ci ne s'affichent pas correctement. Pour obtenir le bon affichage, à partir d'Excel allez dans « Fichier », « Ouvrir », dans le menu déroulant choisissez « Tous les fichiers (*.*) », sélectionnez le fichier à ouvrir, puis cliquez sur « Ouvrir ». Un assistant d'importation de texte s'ouvre, cliquez sur « Délimité » puis « suivant ». Ensuite cliquez sur séparateurs « Points-virgules » puis « Terminer ». Votre fichier s'ouvre alors correctement dans Excel.

Interprétations des fichiers d'analyses de tendances

1) Le graphique de variation d'abondances au cours du temps par espèce :

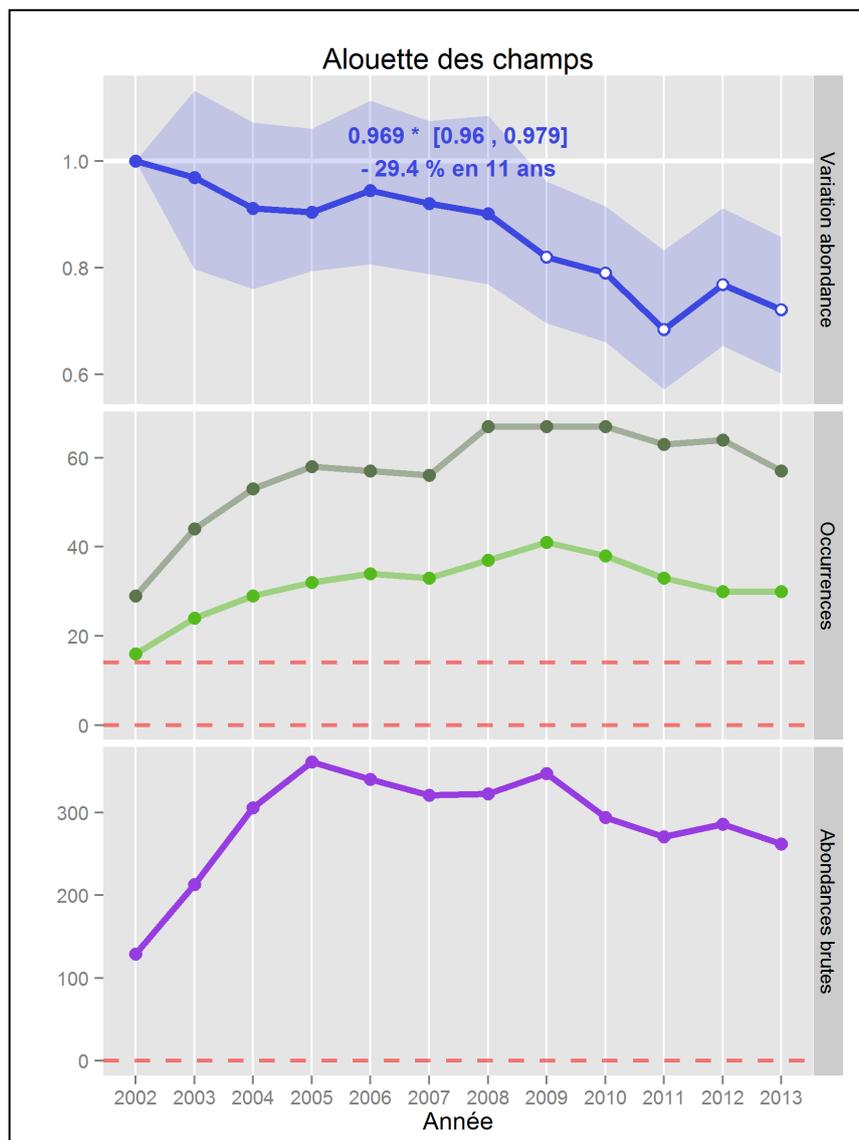
Pour chaque espèce, un graphique est créé. Il se compose de 3 sous représentations graphiques :

- **Les variations d'abondance** : en bleu, représente les variations inter-annuelles des effectifs de l'espèce, à partir d'une modèle linéaire généralisé (glm) de loi quasi-poisson qui modélise les variations d'abondances en fonction du carré et de l'année, ainsi qu'en prenant en compte la sur-dispersion des données :

Glm(effectif)~carré+année

Le coefficient année du modèle est récupéré pour chaque année et permet de tracer le graphe. La première année est fixée arbitrairement à 1 (on peut lire le graphe comme 1^{ère} année équivalente à 100% de l'abondance relative), les coefficients des années suivantes représentent les variations d'abondances par rapport à cette année de référence. La ligne en bleu représente la valeur de variation d'abondance la plus probable, et le halot en bleu clair autour de la courbe est son intervalle de confiance. La première valeur inscrite sur le graphe est la tendance moyenne de variation d'abondance ainsi que son intervalle de confiance. La présence d'un astérisque signifie que cette valeur est significative ($p.value < 0,05$). Dans ce cas le pourcentage de variation d'abondance de l'espèce sur la période considérée est également affiché.

Il y a possibilité d'enregistrer pour chaque espèce uniquement ce graphe de variation d'abondance (voir partie « Lancement du script R» du tutoriel).



- **Les occurrences** : en vert, représente en vert foncé le nombre de carré STOC prospecté par an, et en vert clair le nombre de carré où l'espèce est présente par an. Les lignes rouges représentent des seuils de robustesses : si l'occurrence médiane de l'espèce est inférieure à 14, l'analyse de variations d'abondance de celle-ci est considérée comme incertaine.
- **Les abondances brutes** : en violet, représente le nombre d'individu de l'espèce, tous carrés confondus, par an.

2) Le fichier des variations annuelles par espèce :

	A	B	C	D	E	F	G	H	I	J	K
1	id	code_espece	nom_espece	indicateur	annee	abondance_relative	IC_inferieur	IC_superieur	erreur_standard	p_value	significatif
2	20151020-19H42	CARCAN	Linotte mélodieuse	TRUE	2002	1	1	1	0	1	FALSE
3	20151020-19H42	CARCAN	Linotte mélodieuse	TRUE	2003	1,44	0,88	2,078	0,2497	0,145	FALSE
4	20151020-19H42	CARCAN	Linotte mélodieuse	TRUE	2004	2,059	1,273	3,163	0,337	0,002	TRUE
5	20151020-19H42	CARCAN	Linotte mélodieuse	TRUE	2005	1,594	1,059	2,703	0,4909	0,051	FALSE
6	20151020-19H42	CARCAN	Linotte mélodieuse	TRUE	2006	1,028	0,615	1,803	0,4131	0,915	FALSE
7	20151020-19H42	CARCAN	Linotte mélodieuse	TRUE	2007	0,611	0,358	1,038	0,2914	0,083	FALSE
8	20151020-19H42	CARCAN	Linotte mélodieuse	TRUE	2008	1,135	0,73	1,733	0,1534	0,613	FALSE
9	20151020-19H42	CARCAN	Linotte mélodieuse	TRUE	2009	0,874	0,537	1,367	0,2934	0,602	FALSE
10	20151020-19H42	CARCAN	Linotte mélodieuse	TRUE	2010	0,863	0,576	1,469	0,2295	0,576	FALSE
11	20151020-19H42	CARCAN	Linotte mélodieuse	TRUE	2011	1,303	0,765	2,093	0,222	0,304	FALSE
12	20151020-19H42	CARCAN	Linotte mélodieuse	TRUE	2012	1,136	0,7	1,863	0,3373	0,622	FALSE
13	20151020-19H42	CARCAN	Linotte mélodieuse	TRUE	2013	1,117	0,715	1,765	0,3007	0,675	FALSE
14	20151020-19H42	CARCAR	Chardonneret élégant	TRUE	2002	1	1	1	0	1	FALSE
15	20151020-19H42	CARCAR	Chardonneret élégant	TRUE	2003	0,744	0,61	0,967	0,138	0,032	TRUE
16	20151020-19H42	CARCAR	Chardonneret élégant	TRUE	2004	0,574	0,445	0,737	0,1058	0	TRUE

Ce fichier contient les sorties et ses interprétations du modèle statistique (décrit dans la partie 1) de cette section) qui permet de voir les variations inter-annuelles d'abondance des populations d'oiseaux pour chaque espèce. C'est grâce à ses valeurs que les représentations graphiques sont réalisées. Par ordre d'apparition dans le fichier, les colonnes sont :

- **id** : identifiant de l'analyse, par défaut au format date-heure, ou avec le nom de votre choix
- **code_espece** : code à 6 lettres de l'espèce
- **nom_espece** : nom vernaculaire de l'espèce
- **indicateur** : si l'espèce est une espèce spécialiste d'un habitat donné ou généraliste, et donc utilisée dans le calcul de l'indicateur par groupe de spécialistes
- **annee** : année associée à la valeur de l'abondance relative estimée par le modèle
- **abondance_relative** : abondance relative annuelle estimée par le modèle. La première année est fixée à 1
- **IC_inferieur** : intervalle de confiance inférieur de l'abondance relative estimée
- **IC_superieur** : intervalle de confiance supérieur de l'abondance relative estimée
- **erreur_standard** : erreur standard annuelle associée à l'abondance relative
- **p_value** : valeur qui permet de savoir si l'abondance relative est significative (significative si $p.value < 0,05$). Valeur arrondie à 3 chiffres après la virgule
- **significatif** : si la valeur d'abondance relative est significativement différente de celle de l'année de la première année (année de référence)

- **nb_carre_presence** : nombre de carrés où l'espèce a été contactée l'année considérée
- **abondance** : nombre d'individus contactés l'année considérée

3) Le fichier de la tendance globale par espèce :

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	code_espece	nom_espece	indicateur	nombre_an	premiere_an	derniere_an	tendance	IC_inf	IC_sup	pourcentage_variation	erreur_standard	p_value	significatif	categorie_tendance	mediane_occurrence	valide	raison_incertitude
2	ACCNIS	Epervier d'Europe	FALSE	11	2002	2013	0,943	0,887	1	-47,595	0,03	0,07	FALSE	Incertain	4	Incert	espece trop rare
3	ALAARV	Alouette des champs	TRUE	11	2002	2013	0,969	0,959	0,978	-29,397	0,0048	0	TRUE	Déclin modéré	32,5	bon	
4	ANAPLA	Canard colvert	FALSE	11	2002	2013	1,001	0,974	1,031	1,264	0,0156	0,94	FALSE	Stable	20	bon	
5	ANTPRA	Pipit farlouse	TRUE	11	2002	2013	0,863	0,821	0,912	-80,293	0,0262	0	TRUE	Fort déclin	4	Incert	espece trop rare
6	ANTTRI	Pipit des arbres	FALSE	11	2002	2013	1,028	1,013	1,048	35,23	0,0115	0,01	TRUE	Augmentation modérée	28,5	bon	

Ce fichier contient les sorties et ses interprétations du modèle statistique qui permet d'avoir la tendance globale de la variation d'abondance sur l'ensemble de la période de temps de l'espèce considérée. Le modèle utilisé est le même que pour les variations inter-annuelles, mais utilise les années comme une variable continue. Par ordre d'apparition dans le fichier, les colonnes sont :

- **id** : identifiant de l'analyse, par défaut au format date-heure, ou avec le nom de votre choix
- **code_espece** : code à 6 lettres de l'espèce
- **nom_espece** : nom vernaculaire de l'espèce
- **indicateur** : si l'espèce est une espèce spécialiste d'un habitat donné ou généraliste, et donc utilisée dans le calcul de l'indicateur par groupe de spécialistes
- **nombre_annee** : nombre de pas de temps dans la période considérée
- **premiere_annee** : première année de l'analyse
- **derniere_annee** : dernière année de l'analyse
- **tendance** : estimation par le modèle de la variation moyenne d'abondance sur l'ensemble de la période. Si la valeur est de 1, il n'y a pas de tendance. Une valeur de 0,95 peut donc se lire comme une baisse en moyenne de $1-0,95=0,05$, soit une baisse moyenne de la population de 5% par an.
- **IC_inferieur** : intervalle de confiance inférieur de la tendance estimée
- **IC_superieur** : intervalle de confiance supérieur de la tendance estimée
- **pourcentage_variation** : pourcentage de variation moyenne de l'abondance de la population sur l'ensemble de la période
- **erreur_standard** : erreur standard associée à la tendance estimée
- **p_value** : valeur qui permet de savoir si la tendance est significative (significative si $p.value < 0,05$). Valeur arrondie à 3 chiffres après la virgule
- **significatif** : si la valeur de la tendance est significativement différente d'une tendance stable
- **categorie_tendance_EBCC** : donne la catégorie de variation d'abondance dans laquelle se trouve l'espèce, selon la classification de l'European Bird Census Council. 6 catégories sont possibles : forte augmentation, augmentation modérée, stable, incertain, déclin modéré et fort déclin.

- **mediane_occurrence** : médiane du nombre de fois où l'espèce est contactée par carré et par an. Une médiane de 14 est un seuil de robustesse utilisé dans l'analyse
- **valide** : si l'occurrence médiane de l'espèce est supérieure ou égale à 14, l'espèce est considérée comme suffisamment robuste pour l'analyse par groupe de spécialistes, et sera marquée « bon » dans la colonne. Sinon un poids lui est attribué en fonction de la distribution de ses effectifs, et apparaît en « incertain » dans la colonne
- **raison_incertaine** : donne la raison du classement en « incertain » dans la colonne « valide »

4) Le fichier des variations annuelles par groupe de spécialistes :

Ce fichier contient les variations inter-annuelles d'abondance par groupe de spécialistes par habitats (spécialistes des milieux agricoles, forestiers, bâtis, ouvert et les généralistes). Un indice annuel correspond à la moyenne géométrique des indices annuels des espèces spécialistes du groupe. Par ordre d'apparition dans le fichier, les colonnes sont :

- **groupe** : appartenance du groupe de spécialisation habitat
- **annee** : année associée à la valeur de l'abondance relative estimée
- **abondance_relative** : abondance relative annuelle estimée. La première année est fixée à 1
- **IC_inferieur** : intervalle de confiance inférieur de l'abondance relative estimée
- **IC_superieur** : intervalle de confiance supérieur de l'abondance relative estimée
- **nombre_especes_incertaines** : nombre d'espèces trop rares pour avoir un poids entier dans l'analyse
- **nombre_especes_bonnes** : nombre d'espèces dont la distribution est assez élevée pour avoir un poids entier dans l'analyse

	B	C	D	E	F	G	H
	groupe	annee	abondance_rel	IC_inf	IC_sup	nombre_especes_incertaines	nombre_espece_bonnes
	generaliste	2002	1	1	1	0	14
	generaliste	2003	1,026	0,8	1,304	0	14
	generaliste	2004	0,983	0,772	1,234	0	14
	generaliste	2005	1,056	0,841	1,317	0	14
	generaliste	2006	1,112	0,879	1,394	0	14
	generaliste	2007	1,102	0,862	1,368	0	14
	generaliste	2008	1,067	0,847	1,335	0	14
	generaliste	2009	1,027	0,821	1,304	0	14
	generaliste	2010	1,031	0,821	1,292	0	14
	milieux agricoles	2002	1	1	1	8	12
	milieux agricoles	2003	0,955	0,62	1,471	8	12
	milieux agricoles	2004	1,05	0,701	1,55	8	12
	milieux agricoles	2005	1,033	0,697	1,543	8	12

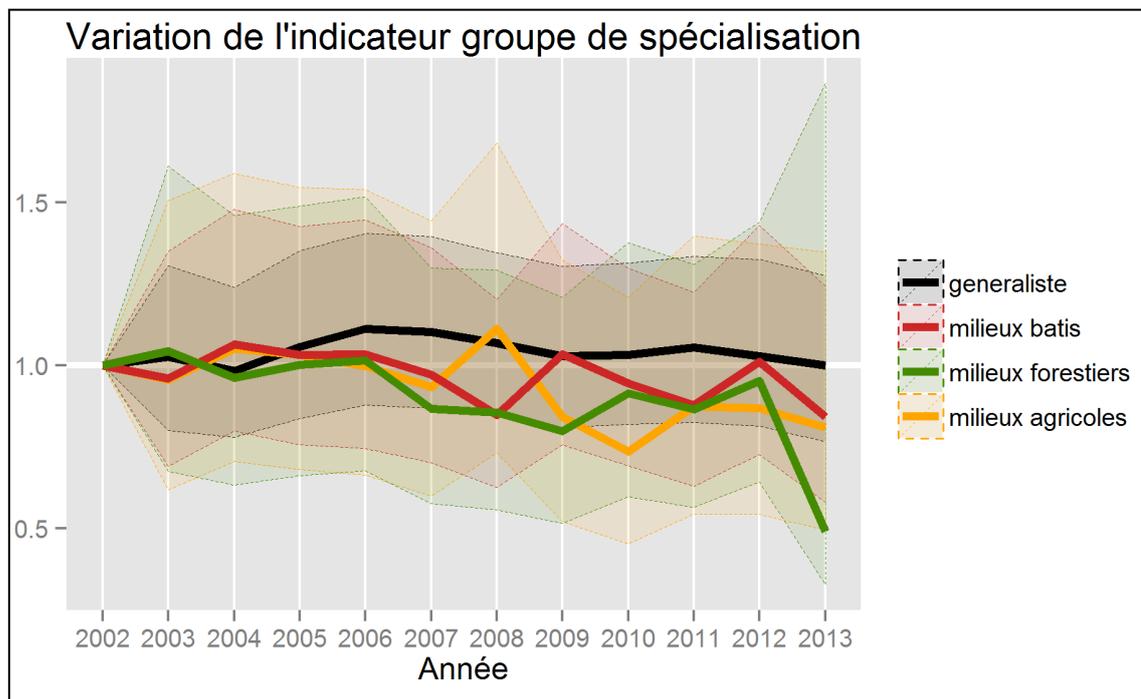
5) Le fichier de la tendance globale par groupe de spécialistes :

	A	B	C	D	E	F
1	id	groupe	tendance	pourcentage_variation	nombre_especes_incertaines	nombre_especes_bonnes
2	20151022-14H19	generaliste	0	0,403	0	14
3	20151022-14H19	milieux agricoles	-0,021	-24,68	8	12
4	20151022-14H19	milieux batis	-0,011	-13,196	3	10
5	20151022-14H19	milieux forestiers	-0,029	-34,498	11	11
6						

Ce fichier contient les sorties du modèle statistique qui permet d'avoir la tendance globale de la variation d'abondance sur l'ensemble de la période de temps des groupes de spécialistes. Le modèle utilisé est une régression linéaire des abondances relatives en fonction du temps (voir partie 4) de cette section). Par ordre d'apparition dans le fichier, les colonnes sont :

- **id** : identifiant de l'analyse, par défaut au format date-heure, ou avec le nom de votre choix
- **groupe** : appartenance du groupe de spécialisation habitat
- **tendance** : estimation par le modèle de la variation moyenne d'abondance par pas de temps
- **pourcentage_variation** : pourcentage de variation moyenne de l'abondance du groupe de spécialistes sur l'ensemble de la période
- **nombre_especes_incertaines** : nombre d'espèces trop rares pour avoir un poids entier dans l'analyse
- **nombre_especes_bonnes** : nombre d'espèces dont la distribution est assez élevée pour avoir un poids entier dans l'analyse

6) **Le graphique de variation d'abondances au cours du temps par groupe de spécialistes :**



Ce graphe représente les variations inter-annuelles d'abondance par groupe de spécialistes. La première année est fixée arbitrairement à 1 (on peut lire le graphe comme 1^{ère} année équivalente à 100% de l'abondance relative), les coefficients des années suivantes représentent les variations d'abondances par rapport à cette année de référence. Les lignes nettes épaisses représentent les valeurs de variations d'abondances

les plus probables, et les halots en clair autour de ces courbes sont leurs intervalles de confiance de couleurs respectives.

7) Le fichier d'informations relatives au calcul de l'indicateur par groupe de spécialistes :

	A	B	C	D	E	F	G	H	I	J	K	L
1	code_espece	nom_espece	annee	groupe_indicat	poids_erreur_standard	poids_incertitude	poids_final	abondance	IC_inferieur	IC_superieur	valide	mediane_occurrence
2	ACCNIS	Epervier d'Euro	2003	non	1	1	1	1	1	1	bon	15
3	ACCNIS	Epervier d'Euro	2004	non	1	1	1	1,934	1,033	3,009	bon	15
4	ACCNIS	Epervier d'Euro	2005	non	1	1	1	1,319	0,746	2,124	bon	15
5	ACCNIS	Epervier d'Euro	2006	non	1	1	1	2,053	1,198	3,262	bon	15
6	ACCNIS	Epervier d'Euro	2007	non	1	1	1	1,496	0,861	2,316	bon	15
7	ACCNIS	Epervier d'Euro	2008	non	1	1	1	1,942	1,048	3,473	bon	15
8	ACCNIS	Epervier d'Euro	2009	non	1	1	1	1,55	0,856	2,539	bon	15
9	ACCNIS	Epervier d'Euro	2010	non	1	1	1	1,168	0,609	2,058	bon	15
10	ACCNIS	Epervier d'Euro	2011	non	1	1	1	1,576	0,897	2,669	bon	15
11	ACCNIS	Epervier d'Euro	2012	non	1	1	1	1,533	0,85	2,593	bon	15
12	ACCNIS	Epervier d'Euro	2013	non	1	1	1	1,767	0,966	3,005	bon	15
13	ANTPRA	Pipit farlouse	2003	milieux agricole	1	0,84	0,84	1	1	1	Incerta	9
14	ANTPRA	Pipit farlouse	2004	milieux agricole	1	0,84	0,84	0,883	0,671	1,122	Incerta	9
15	ANTPRA	Pipit farlouse	2005	milieux agricole	1	0,84	0,84	1,006	0,744	1,254	Incerta	9
16	ANTPRA	Pipit farlouse	2006	milieux agricole	1	0,84	0,84	0,29	0,164	0,445	Incerta	9

Ce fichier contient le détail des poids attribués à chaque espèce et chaque année dans le calcul de l'indicateur par groupe de spécialistes. En effet, l'effectif de certaines espèces étant trop faible pour avoir une bonne estimation de leur abondance relative, un poids leur est attribué. Par ordre d'apparition dans le fichier, les colonnes sont :

- **code_espece** : code à 6 lettres de l'espèce
- **nom_espece** : nom vernaculaire de l'espèce
- **annee** : année associée à la valeur de l'abondance relative estimée par le modèle
- **groupe_indicateur** : non de la spécialité si l'espèce est une espèce spécialiste d'un habitat donné ou généraliste, et donc utilisée dans le calcul de l'indicateur par groupe de spécialistes
- **poids_erreur_standard** : valeur du poids attribuée à l'abondance relative annuelle en fonction de l'erreur standard du modèle et des intervalles de confiance (0 ou 1)
- **poids_incertitude** : valeur du poids attribuée à l'abondance relative annuelle en fonction de l'abondance de l'espèce (poids maximal = 1)
- **poids_final** : poids final attribué à l'abondance relative annuelle de l'espèce dans le calcul de l'indicateur par groupe de spécialistes. Cette valeur est la multiplication du poids de l'erreur standard et du poids liée à l'incertitude de l'estimation
- **abondance_relative** : abondance relative annuelle estimée par le modèle. La première année est fixée à 1
- **IC_inferieur** : intervalle de confiance inférieur de l'abondance relative estimée
- **IC_superieur** : intervalle de confiance supérieur de l'abondance relative estimée
- **valide** : si l'occurrence médiane de l'espèce est supérieure ou égale à 14, l'espèce est considérée comme suffisamment robuste pour l'analyse par groupe de spécialistes, et sera marquée « bon » dans la colonne. Sinon un poids lui est attribué en fonction de la distribution de ses effectifs, et apparaît en « incertain » dans la colonne

- **mediane_occurence** : médiane du nombre de fois ou l'espèce est contactée par carré et par an. Une médiane de 14 est un seuil de robustesse utilisé dans l'analyse



Vous avez toutes les clefs en mains, maintenant c'est à vous !

N'hésitez pas à nous poser des questions pour tous problèmes rencontrés tant sur le lancement des analyses dans R que sur l'interprétation des résultats :

dgonzalez@mnhn.fr

fjiguet@mnhn.fr

VIGIENATURE
Un réseau de citoyens qui fait avancer la science