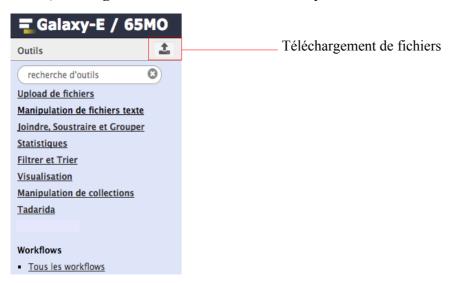




<u>Test analyse de données du programme STOCeps de vigie</u> nature sur la plateforme Galaxy-E

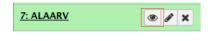
1) Chargement des fichiers dans Galaxy-E



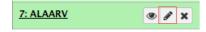
Pour importer les fichiers, il faut cliquer sur le bouton téléchargement en haut à gauche. Il est préférable de les importer au format **tsv** ou **csv** ainsi qu'en **utf8**.



La fenêtre « **Téléverser depuis le web ou à partir de votre pc** » s'affiche. Il faut sélectionner « **choisir un fichier en local** » puis « **démarrer** » pour charger les données. Une fois que les données sont chargées, la fenêtre peut être fermée. Dans type, il est préférable de sélectionner **tabular**. Les données s'affichent à droite une fois chargées.



Pour vérifier qu'il n'y a pas eu de problèmes lors de l'importation, il faut cliquer sur l'œil à droite des noms de jeux de données.



Pour modifier le nom du tableau ou ajouter des informations sur la table, il faut cliquer sur le crayon à droite des noms de jeux de données.

La rechercher d'outils peut être effectué grâce à la barre de recherche en haut à gauche (« recherche d'outils »).

2) Obtenir une base de données ALAARV

Grâce à L'outil "Filtrer des données sur une colonne en utilisant des expressions simples".

Il permet de créer une fichier tabulé avec les données pour une espèce (Exemple ici pour ALAARV).

Pour sélectionner que les lignes mentionnant l'espèce ALAARV dans la troisième colonne du jeu de données, il faut utiliser la condition suivante : c3=='ALAARV'.



3) Obtenir le nombre de carrés STOC par année pour l'espèce ALAARV et pour toutes les espèces

Grâce à l'outil « Compter le nombre d'occurrences de chaque enregistrement ».

Il compte les occurrences de valeurs uniques dans la(les) colonne(s) sélectionnée(s)".

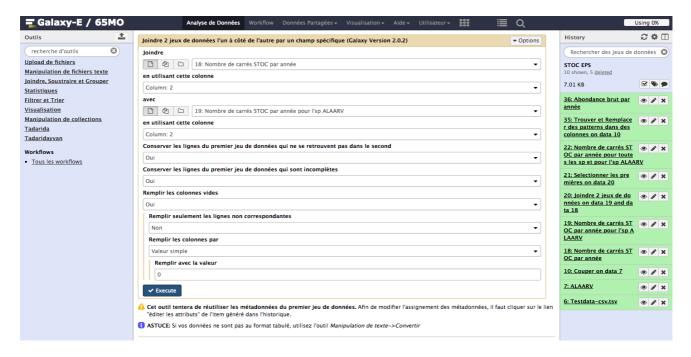
Il permet d'obtenir le nombre de carrés STOC par année pour toutes les espèces ou pour une espèce (ex : ALAARV). Il suffit d'indiquer le champ dans lequel on veut compter les termes dans « Select/Unselect all ».



4) Joindre le nombre de carrés STOC par année pour toutes les espèces et pour l'espèce ALAARV en une table

Grâce à l'outil « Joindre les lignes de deux jeux de données l'un à côté de l'autre par un champ spécifique ».

Cet outil joint les lignes de deux jeux de données (ici le nombre de carrés STOC par année pour toutes les espèces avec le nombre de carrés STOC par année pour l'espèce ALAARV) pour un champ spécifique (ici le champ année).



5) Trier les colonnes

Grâce à l'outil « Couper des colonnes d'un jeu de données tabulé ».

Il permet de sélectionner des colonnes spécifiques à partir d'un jeu de données. Il a permis d'enlever la colonne carré ou encore de trier les colonnes dans la base de données nombre de carrés STOC par année.

Il faut indiquer dans « Couper les colonnes », quelles colonnes on veut afficher ainsi que l'ordre.



6) Convertir le champ année en commentaire

Grâce à l'outil « <u>Trouver et Remplacer des patterns dans des colonnes en utilisant des expressions</u> régulières (regex) ».

Cet outil travaille ligne après ligne sur la donnée spécifiée en entrée et remplace le texte correspondant aux patterns d'expression régulière rentré par la correspondance proposée. Cet outil utilise les expressions régulières du langage de programmation **Python**.

Il a permis de convertir le champ année en commentaire pour que la somme (présenté partie 7) soit effectuée uniquement sur le nombre d'individus.



7) Additionner les abondances pour une année

Grâce à l'outil « Grouper des données par une colonne et pratiquer des opérations d'agrégation sur d'autres colonnes ».

Il permet de grouper les jeux de données d'entrée par une colonne particulière et d'appliquer des fonctions d'**agrégation** (ici une somme entre les abondances pour une année).





8) Obtenir les fichiers des variations annuelles par espèce et de la tendance globale par espèce

Grâce à l'outil : « <u>STOCeps Création des fichiers des variations annuelles par espèces et de la tendance globale par espèce</u> ».

L'outil doit être exécuté avec le jeu de données initiale.



Il permet d'obtenir les fichiers :

• variations annuelles par espèce

Ce fichier contient les sorties et les interprétations du modèle statistique qui permet de voir les variations inter-annuelles d'abondance des populations d'oiseaux pour chaque espèce. Avec ce jeux de données, l'outil "Filtrer des données sur une colonne en utilisant des expressions simples" permet de sélectionner les informations sur l'espèce voulu (ici l'espèce ALAARV) pour créer le graphique "Variation d'abondance pour l'espèce ALAARV". Il faut indiquer la condition : c2=='ALAARV'.

2	3	4	5	6	7	8	9	10	11	12	13	14
code_espece	nom_espece	indicateur	annee	abondance_relative	IC_inferieur	IC_superieur	erreur_standard	p_value	significatif	nb_carre	nb_carre_presence	abondance
ACCNIS	Epervier d'Europe	non	2001	1	1	1	0	1	FALSE	15	4	4
ACCNIS	Epervier d'Europe	non	2002	0.216	0.092	0.694	0.547	0.005	TRUE	46	4	4
ACCNIS	Epervier d'Europe	non	2003	0.464	0.225	1.023	0.0989	0.094	FALSE	55	6	7
ACCNIS	Epervier d'Europe	non	2004	0.831	0.395	1.791	0.1924	0.655	FALSE	57	9	11
ACCNIS	Epervier d'Europe	non	2005	0.343	0.154	1.26	0.435	0.042	TRUE	55	4	4
ACCNIS	Epervier d'Europe	non	2006	0.658	0.271	1.374	0.1508	0.341	FALSE	56	7	7
ACCNIS	Epervier d'Europe	non	2007	0.278	0.093	0.67	0.3479	0.016	TRUE	52	3	3
ACCNIS	Epervier d'Europe	non	2008	0.157	0.03	0.623	0.2101	0.015	TRUE	44	1	1
ACCNIS	Epervier d'Europe	non	2009	0.416	0.189	1.302	0.0777	0.077	FALSE	48	4	4
ACCNIS	Epervier d'Europe	non	2010	0.859	0.396	2.367	0.182	0.728	FALSE	43	4	8
ACCNIS	Epervier d'Europe	non	2011	0.389	0.164	1.007	0.5147	0.116	FALSE	27	2	2
ACCNIS	Epervier d'Europe	non	2012	1.646	0.477	6.856	0.2533	0.444	FALSE	23	3	3
ACCNIS	Epervier d'Europe	non	2013	0.504	0.161	1.549	0.9958	0.258	FALSE	30	3	3
ACCNIS	Epervier d'Europe	non	2014	5.277	0.883	31.522	0.5052	0.098	FALSE	7	1	1
ACRSCI	Rousserolle effarvate	non	2001	1	1	1	0	1	FALSE	15	1	4
ACRSCI	Rousserolle effarvate	non	2002	0	0	NA	1886.6398	0.992	FALSE	46	0	0
ACRSCI	Rousserolle effarvate	non	2003	0.29	0.196	0.475	0	0	TRUE	55	1	2
ACRSCI	Rousserolle effarvate	non	2004	0	0	NA	429.9583	0.989	FALSE	57	0	0
ACRSCI	Rousserolle effarvate	non	2005	0.237	0.154	0.347	0	0	TRUE	55	2	2

• tendance globale par espèce

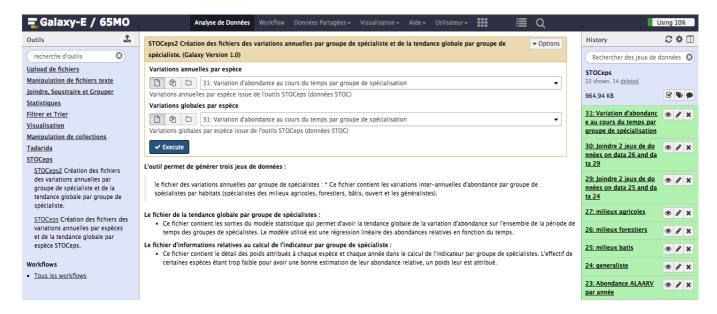
Ce fichier contient les sorties et ses interprétations du modèle statistique qui permet d'avoir la tendance globale de la variation d'abondance sur l'ensemble de la période de temps de l'espèce considérée.

2	3	4	5	6	7	8	9	10	11	12	13	14
code_espece	nom_espece	indicateur	nombre_annees	premiere_annee	derniere_annee	tendance	IC_inferieur	IC_superieur	pourcentage_variation	erreur_standard	p_value	significatif
ACCNIS	Epervier d'Europe	non	13	2001	2014	1.021	0.959	1.089	30.806	0.0317	0.507	FALSE
ACRSCI	Rousserolle effarvate	non	13	2001	2014	0.983	0.953	1.025	-19.63	0.0187	0.377	FALSE
AEGCAU	Mésange à longue queue	non	13	2001	2014	1.041	0.992	1.086	69.206	0.024	0.08	FALSE
ALAARV	Alouette des champs	milieux agricoles	13	2001	2014	0.976	0.965	0.993	-26.71	0.0081	0.004	TRUE
ANAPLA	Canard colvert	non	13	2001	2014	1.054	1.02	1.083	97.781	0.0196	0.005	TRUE
ANTTRI	Pipit des arbres	non	13	2001	2014	0.996	0.977	1.019	-5.54	0.012	0.715	FALSE
APUAPU	Martinet noir	milieux batis	13	2001	2014	0.95	0.921	0.982	-48.396	0.0161	0.003	TRUE
ARDCIN	Héron cendré	non	13	2001	2014	0.985	0.95	1.026	-18.168	0.019	0.426	FALSE
BUTBUT	Buse variable	milieux agricoles	13	2001	2014	1.037	0.999	1.072	60.334	0.0186	0.044	TRUE
CARCAN	Linotte mélodieuse	milieux agricoles	13	2001	2014	0.837	0.805	0.866	-90.102	0.0188	0	TRUE
CARCAR	Chardonneret élégant	milieux batis	13	2001	2014	0.888	0.855	0.915	-78.696	0.0159	0	TRUE
CARCHL	Verdier d'Europe	milieux batis	13	2001	2014	0.904	0.885	0.918	-73.232	0.0082	0	TRUE
CERBRA	Grimpereau des jardins	milieux forestiers	13	2001	2014	0.974	0.954	1	-28.558	0.0127	0.048	TRUE
CETCET	Bouscarle de Cetti	non	13	2001	2014	0.956	0.922	0.99	-44.206	0.0191	0.025	TRUE
CIRAER	Busard des roseaux	non	13	2001	2014	0.821	0.783	0.855	-92.311	0.0249	0	TRUE
COLLIV	Pigeon biset	non	13	2001	2014	1.014	0.967	1.061	19.849	0.0275	0.607	FALSE
COLPAL	Pigeon ramier	generaliste	13	2001	2014	1.033	1.022	1.042	53.411	0.0049	0	TRUE
CORCOR	Corneille noire	generaliste	13	2001	2014	0.976	0.965	0.986	-26.777	0.0072	0.001	TRUE
CORMON	Choucas des tours	milieux batis	13	2001	2014	1.053	1.011	1.095	96.23	0.0202	0.007	TRUE
CUCCAN	Coucou gris	generaliste	13	2001	2014	1.022	1	1.039	33.332	0.01	0.024	TRUE
DENMAJ	Pic épeiche	milieux forestiers	13	2001	2014	1.006	0.978	1.033	8.183	0.0144	0.673	FALSE

9) Obtenir les fichiers des variations annuelles par groupe de spécialistes, de la tendance globale par groupe de spécialistes et des informations relatives au calcul de l'indicateur par groupe de spécialistes

Grâce à l'outil : « STOCeps2 Création des fichiers des variations annuelles par groupe de spécialiste et de la tendance globale par groupe de spécialiste ».

L'outil doit être exécuté avec les deux jeux de données : « variations annuelles par espèce » et « tendance globale par espèce ».



Il permet d'obtenir les fichiers :

· variations annuelles par groupe de spécialistes

Ce fichier contient les variations inter-annuelles d'abondance par groupe de spécialistes par habitats. Un indice annuel correspond à la moyenne géométrique des indices annuels des espèces spécialistes du groupe.

1	2	3	4	5	6	7
groupe	annee	abondance_relative	IC_inferieur	IC_superieur	nombre_especes_incertaines	nombre_espece_bonnes
generaliste	2001	1	1	1	1	12
generaliste	2002	1.151	0.848	1.594	1	12
generaliste	2003	1.056	0.769	1.473	1	12
generaliste	2004	1.057	0.761	1.484	1	12
generaliste	2005	1.207	0.888	1.665	1	12
generaliste	2006	1.312	0.969	1.798	1	12
generaliste	2007	1.279	0.919	1.758	1	12
generaliste	2008	1.257	0.908	1.765	1	12
generaliste	2009	1.142	0.817	1.579	1	12
generaliste	2010	1.28	0.919	1.829	1	12
generaliste	2011	1.192	0.844	1.674	1	12
generaliste	2012	1.304	0.932	1.882	1	12
generaliste	2013	1.198	0.841	1.706	1	12
generaliste	2014	1.127	0.706	1.845	1	12
milieux agricoles	2001	1	1	1	4	7
milieux agricoles	2002	0.978	0.501	1.921	4	7
milieux agricoles	2003	0.829	0.374	1.763	4	7
milieux agricoles	2004	0.845	0.384	1.78	4	7
milieux agricoles	2005	1.138	0.527	2.367	4	7
milieux agricoles	2006	0.874	0.446	1.737	4	7

Avec ce jeu de données, l'outil "Filtrer des données sur une colonne en utilisant des expressions simples" permet de sélectionner les informations sur le groupe de spécialistes voulu (généralistes, milieux agricoles, milieux bâtis et milieux forestiers).

Après cela, il faut utiliser l'outil « <u>Joindre les lignes de deux jeux de données l'un à côté de l'autre</u> <u>par un champ spécifique</u> » pour joindre côte à côte les informations de variation d'abondance sur les groupes. Ce jeu de données permet de créer le graphique "Variation d'abondances au cours du temps par groupe de spécialistes".

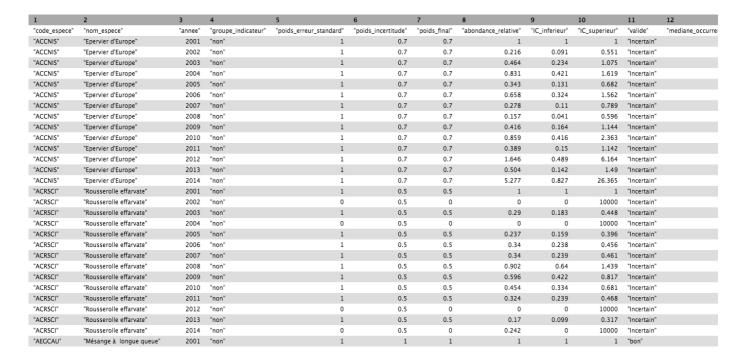
• tendance globale par groupe de spécialistes

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.

1	2	3	4	5	6
id	groupe	tendance	pourcentage_variation	Incertain	bon
20170818-10H41	generaliste	0.011	14.794	1	12
20170818-10H41	milieux agricoles	-0.008	-10.071	4	7
20170818-10H41	milieux batis	-0.038	-49.289	3	7
20170818-10H41	milieux forestiers	-0.027	-35.671	5	8

informations relatives au calcul de l'indicateur par groupe de spécialistes

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.



9) Création de graphiques

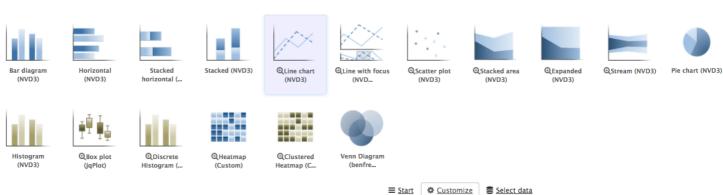
Il faut cliquer sur « visualisation » et sur « charts » \rightarrow



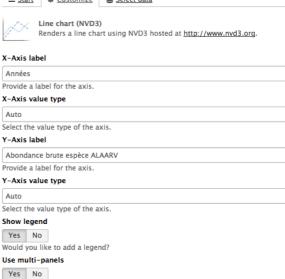
Pour obtenir les graphique :

- du nombre de carrés STOC par année pour toutes les espèces et pour l'espèce ALAARV
- de <u>l'abondance brute d'individus de l'espèce ALAARV tous carrés confondus par an</u>
- de la <u>variation d'abondance pour l'espèce ALAARV</u>
- de la variation d'abondances au cours du temps par groupe de spécialistes

Sélectionner le type de graphique :



Ajouter un titre et des noms aux axes →



Would you like to separate your data into individual panels?

1: Data series

Provide a label

Abondance brute ALAARV

Pick a series color

■ Select a color

Data point labels

Column: Row Number

Values for x-axis

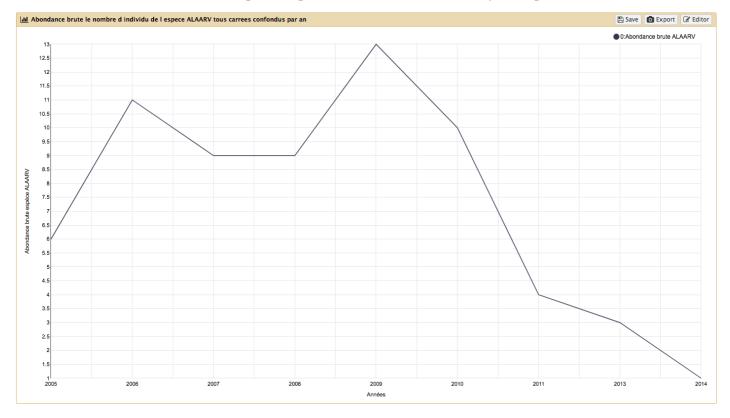
Column: 1

Values for y-axis

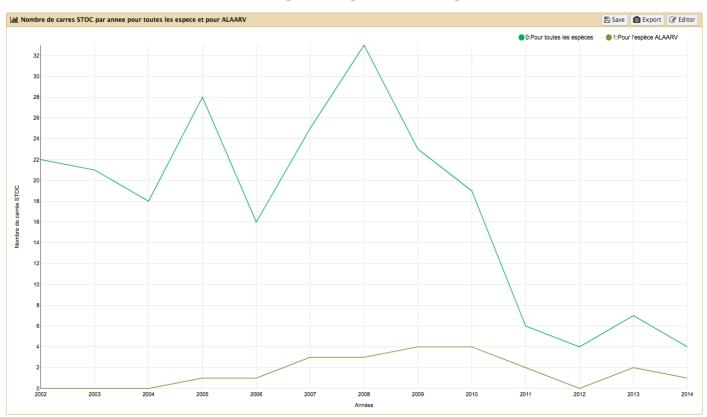
Column: 2

10) Les graphiques

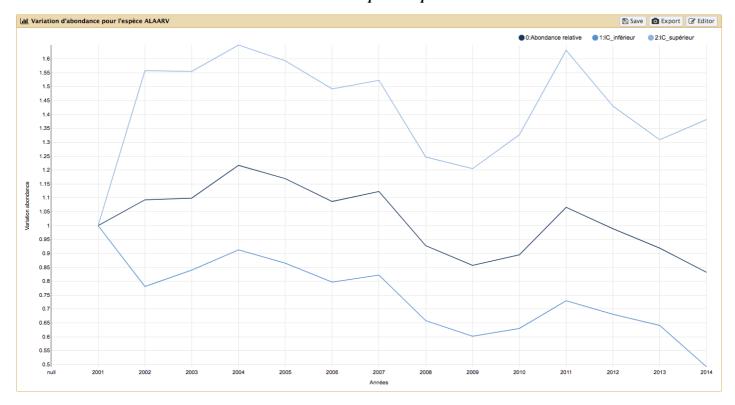
Abondance brute pour l'espèce ALAARV tous carrées confondus par an



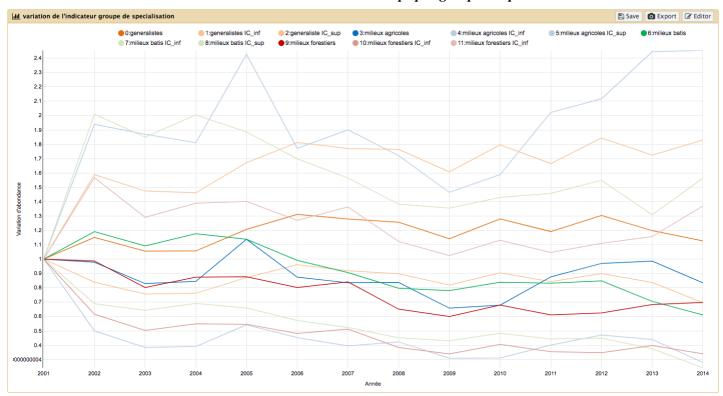
Nombre de carrées STOC par année pour toutes les espèces ALAARV



Variation d'abondance pour l'espèce ALAARV



Variation d'abondances au cours du temps par groupe de spécialistes

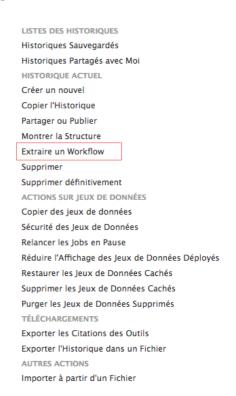


11) Création du workflow

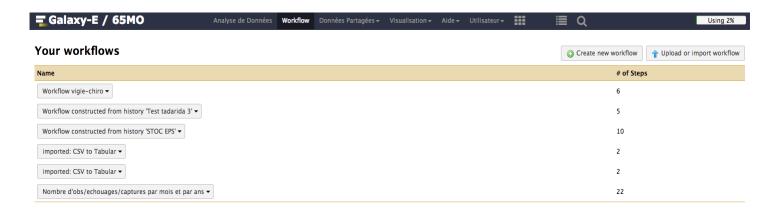
Sur Galaxy-E, il est possible de créer un **workflow** pour automatiser toutes les étapes de traitement des bases de données. Un workflow est la description d'une suite de tâches permettant un enchaînement automatisé de différentes opérations et étapes. Sur Galaxy-E, il est possible de créer un workflow pour automatiser toutes les étapes de traitement des jeux de données. Quand ce workflow est lancé avec un jeu de données STOCeps, il permet d'obtenir les jeux de données servants à la création des graphiques.

Pour extraire un workflow, il faut cliquer sur les « options de l'historique » puis « extraire un Workflow ».





Dans « **workflow** », il est possible de modifier manuellement son workflow ou de l'exécuter sur un jeu de données/fichier audio en appuyant sur « **run** ».



Workfow final

