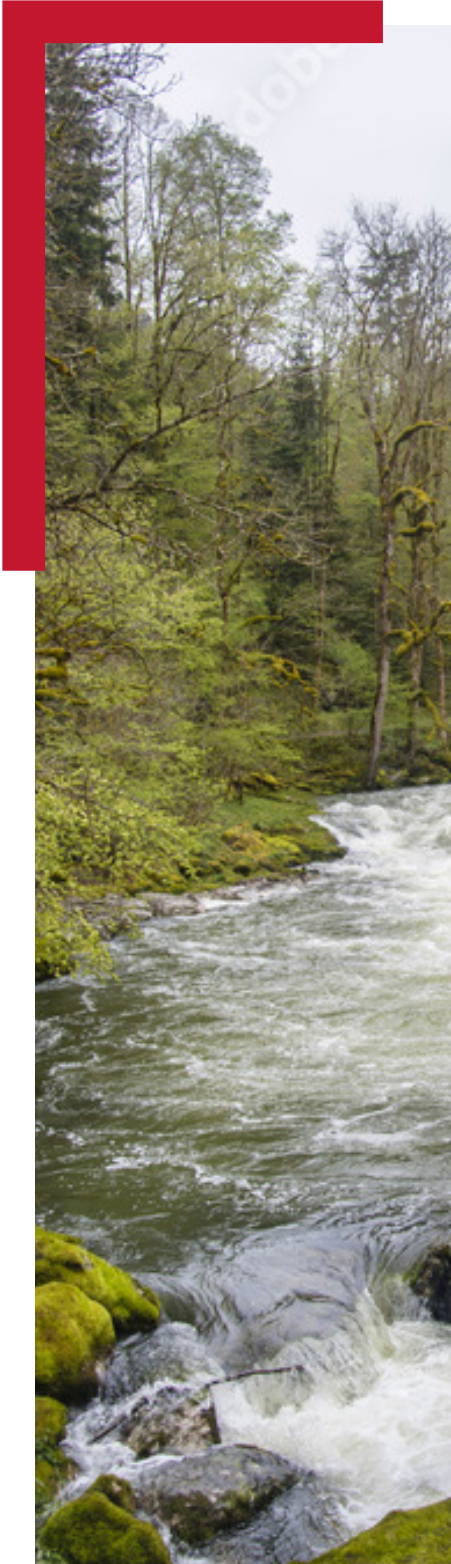




CERTIFICAT EN ANALYSE DE DONNÉES POUR L'ÉCOLOGIE, LE SUIVI DE LA BIODIVERSITÉ ET LA GESTION DES ESPACES NATURELS



L'École Pratique des Hautes Études vous propose une formation en trois niveaux aux méthodes généralistes d'analyse de données appliquées à l'écologie. Ce parcours vous donnera les compétences fondamentales nécessaires à comprendre des résultats d'analyse statistique, puis à construire, analyser et communiquer vous-mêmes vos propres jeux de données écologiques. Vous aborderez les notions fondamentales et les mettrez en application dans un esprit pratique, à partir de véritables cas d'étude.

Vous deviendrez progressivement autonome, dans la mesure de vos besoins : comprendre l'essentiel, mettre en pratique, puis devenir force de proposition. Vous vous préparerez à aborder des méthodes plus spécialisées : modélisation d'abondances ou de dynamiques de populations, modèles de distributions d'espèces, analyses génétiques ou acoustiques...

La formation est structurée en trois niveaux. Le niveau de découverte est accessible à tous sans prérequis initial. Un test de positionnement vous permet d'évaluer vos compétences préalables si vous souhaitez démarrer aux niveaux suivants.

INFORMATIONS

Formateur : Jean-Yves Barnagaud (jean-yves.barnagaud@ephe.psl.eu)

Lieu : Maison des Sciences de l'Homme, 54 boulevard Raspail 75006 Paris.
Métro Sèvres Babylone ou Rennes.

Format présentiel uniquement

Dates en 2025 : 22 - 28 mars

Horaires : 9h30 - 12h30, 13h30 - 17h (sauf lundi : début à 10h, vendredi : fin à 15h30)

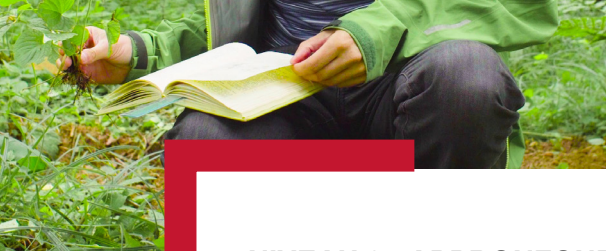
Matériel à prévoir :

- Un ordinateur portable avec MS Excel ou tout autre logiciel de tableur.
- Installer au préalable les logiciels gratuits R et R-studio, téléchargeables à ces adresses (contacter le formateur en cas de difficultés) :

<https://cran.r-project.org/>

<https://rstudio.com/products/rstudio/download/>

- Des directives pour l'installation de bibliothèques R complémentaires vous seront communiquées par le formateur une semaine à dix jours avant le début de la formation (droits administrateur nécessaires pour l'installation).



NIVEAU 3 : APPRONFONDISSEMENT

Modélisation statistique de données écologiques : une approche pratique

Cette semaine vise à vous rendre aussi autonome que possible dans votre pratique des analyses de données au quotidien, en vous faisant passer de l'utilisation éclairée de méthodes connues (niveau 2 – renforcement) à **la mobilisation de nouvelles méthodes** répondant à des enjeux émergeant de jeux de données complexes : échantillonnages stratifiés (multiples régions ou espèces), données opportunistes issues des sciences participatives, grandes tables de communautés. A l'issue de cette formation, vous devriez **être capable d'élaborer vous-même des stratégies d'analyse**, de les soumettre à la discussion collective, et de mobiliser de nouvelles méthodes à partir de la littérature.

Si, dans les deux premiers niveaux de ce certificat, nous avons surtout ciblé la compréhension des méthodes, ce troisième module mobilisera surtout votre créativité, votre capacité à croiser les informations et à vous approprier des situations inconnues : vous deviendrez ainsi une véritable **personne-ressource** dans votre structure. Parce qu'il s'agit moins de maîtriser un grand nombre de techniques que de développer une flexibilité thématique et méthodologique, nous n'aborderons pas d'outils statistiques fondamentalement nouveaux par rapport au niveau de renforcement. Modèles linéaires et analyses multivariées resteront notre base, mais vous les utiliserez différemment, en vous adaptant au contexte. L'essentiel du module prend la forme de **mises en situation** issues de véritables collaborations entre porteurs d'enjeux et analystes : estimation des effets du changement climatique sur une espèce protégée, étude des effets de changements paysagers sur une espèce rare, approche fonctionnelle de l'impact d'une espèce sur son milieu, analyse des relations communautés – environnement dans le but d'interpréter des indicateurs de biodiversité...

L'analyse de données écologiques est un travail collaboratif qui rassemble naturalistes, écologues et biostatisticiens. Vous développerez donc **vos positionnement** dans des groupes de travail, en identifiant quels sont vos atouts et vos limites, et comment les mettre à profit dans un raisonnement collectif sur les données. Afin de profiter au mieux de cette approche, nous vous incitons à échanger avec le formateur dès le début de semaine sur les objectifs que vous vous êtes fixé au regard de vos besoins professionnels.

Enfin, une composante importante de cette formation réside dans la **communication** tant des résultats que des méthodes : vous travaillerez votre capacité à expliquer clairement vos choix méthodologiques, en justifier la pertinence, et exposer résultats, limites et perspectives de manière équilibrée et constructive.

Et ensuite ?

En fin de niveau 3, vous disposerez d'une compétence généraliste solide en analyse statistique, qui couvrira l'essentiel de vos besoins quotidiens comme écologue chargé d'analyser des données dans une structure d'expertise ou de gestion-protection des espaces naturels. Vous serez aussi prêt à évoluer vers d'autres formations spécialisées, couvrant des thèmes plus spécifiques (modèles de distributions d'espèces, dynamique de population, estimations d'abondances) ou des outils avancés (inférence bayésienne, science reproductible). Néanmoins, seule la pratique sur vos propres cas d'étude vous permettra d'acquérir l'expérience nécessaire pour devenir véritablement autonome sur les méthodes et leur implémentation. Sachez exploiter l'ensemble des outils disponibles, mais restez critique : les statistiques, en particulier celles adaptées à l'écologie, restent un champ mouvant dans lequel même des méthodes bien rodées sont régulièrement remises en question. Pratiquez, documentez-vous, mettez-vous à l'épreuve, et formez-vous régulièrement à de nouvelles méthodes afin de compléter votre compétence. Sachez aussi partager vos connaissances : c'est un excellent moyen d'évoluer soi-même.

PLANNING

	Heures	Thèmes	Méthodes abordées	Compétences
Lundi	10h 12h	Adapter un modèle aux contraintes de l'échantillonnage : exemple des plans d'échantillonnage	Modèles hiérarchiques et modèles mixtes	<ul style="list-style-type: none"> • Adapter la structure d'un modèle de régression aux hétérogénéités d'un jeu de données structuré • Interpréter les résultats d'un modèle mixte • Identifier les compromis relatifs à la construction d'un modèle statistique sur des données complexes • Construire, implémenter et interpréter un modèle hiérarchique sous R • Justifier ses choix de modélisation et les communiquer
	13h30 17h	Mise en pratique : élaborer, implémenter, contrôler et interpréter un modèle mixte		
Mardi	9h30 12h30	Vers une flexibilisation du cadre de modélisation : introduction au cadre bayésien	Modèles linéaires en cadre bayésien	<ul style="list-style-type: none"> • Comprendre le principe de l'inférence bayésienne et son utilité en analyse de données écologiques • Élaborer un modèle adapté à l'analyse d'un suivi d'espèce rare ou difficilement observable • Explorer la sensibilité d'un modèle aux conditions de paramétrage
	13h30 17h	Pratique des modèles hiérarchiques bayésiens sous R		
Mercredi	9h30 17h	Identifier les enjeux d'analyse et proposer une méthodologie face à un questionnement et un jeu de données inconnus : le cas des données opportunistes en contexte spatial	Analyser la répartition d'un processus écologique dans l'espace à partir de données non protocolées	<ul style="list-style-type: none"> • Élaborer une démarche d'analyse à partir de données inconnues • Comprendre les problèmes posés par les données non protocolées issues d'échantillonnages opportunistes • Développer une stratégie d'analyse créative par une démarche collaborative et itérative
Jeudi	9h30 12h30	Identifier, comprendre et mettre en pratique une méthodologie d'analyse de données en s'appuyant sur la littérature : exemple des données de composition de communautés écologiques	Indices de dissimilarité pour l'écologie, indicateurs de diversité et de composition, analyses multivariées multi-tables	<ul style="list-style-type: none"> • Développer une méthodologie d'analyse face à des données inconnues en s'appuyant sur l'état de l'art • Comprendre les enjeux, caractéristiques et contraintes d'une méthode à partir de la littérature • Elaborer un flux d'analyse dimensionné à la question écologique et aux données disponibles • Argumenter ses choix d'analyse
	13h30 17h30			
Vendredi	9h30 12h30	Questions personnalisées	Tous les thèmes abordés dans la semaine	
	13h30 15h30	Évaluation du niveau 3		