



HAL
open science

Bilan Hcéres de l'unité BioSP. Vague C 2022-2023

Lionel Roques, Denis Allard, Lionel Benoit, Olivier Bonnefon, Jérôme Coville,
Sandy Duperier, Audrey Foncelle, Raphaël Forien, Edith Gabriel, Marie
Grosdidier, et al.

► **To cite this version:**

Lionel Roques, Denis Allard, Lionel Benoit, Olivier Bonnefon, Jérôme Coville, et al.. Bilan Hcéres de l'unité BioSP. Vague C 2022-2023. 2023. hal-04101039

HAL Id: hal-04101039

<https://hal.inrae.fr/hal-04101039>

Submitted on 19 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License

Document d'autoévaluation (DAE) Unité de recherche pluri-équipes

CAMPAGNE D'ÉVALUATION 2022-2023

VAGUE C

L'unité rédige ce document (DAE) en complétant les rubriques ci-dessous et en suivant le plan indiqué. Les parties en bleu, dans le texte et en appendice, sont des repères pour l'autoévaluation. Elles doivent être supprimées dans le document rédigé.

INFORMATIONS GENERALES

Nom de l'unité pour le contrat en cours : Biostatistique et Processus Spatiaux

Acronyme pour le contrat en cours : BioSP

Label et numéro : UR 0546

Nombre d'équipes : 2

Domaine scientifique principal :

ST : Sciences et Technologies

Panels scientifiques (dans la nomenclature du Hcéres) par ordre décroissant d'importance :

Panel 1

ST1 : Mathématiques

Panel 2

Choisissez un élément.

Panel 3

Choisissez un élément.

Panel 4

Choisissez un élément.

Directrice / directeur pour le contrat en cours : Lionel Roques

Établissements et organismes de rattachement (tutelles) :

Liste des établissements et organismes de rattachement (tutelles) de l'unité de recherche **pour le contrat en cours**

- INRAE | - ...
- ... | - ...
- ... | - ...
- ... | - ...
- ... | - ...

1- PRÉSENTATION DE L'UNITÉ

L'unité Biostatistique et processus SPatiaux (BioSP) existe sur le centre INRAE d'Avignon depuis 1975. Elle est historiquement rattachée au département MathNum (mathématiques, informatique, sciences de la donnée et technologies du numérique ; une liste des principaux acronymes utilisés est disponible en annexe à la fin du document) d'INRAE. A sa création, son activité de recherche s'est d'abord concentrée sur les statistiques spatiales avec des intérêts appliqués pour l'environnement physique (sol, climat...). Depuis les années 2000, l'activité scientifique de BioSP s'est diversifiée, notamment par des recrutements de spécialistes des équations aux dérivées partielles (EDP), de numériciens, de probabilistes et de modélisateurs en épidémiologie. Ainsi, l'unité a connu une forte croissance de ses effectifs depuis sa création (**Fig. 1**). En 2018, au moment de la création de la Plateforme nationale d'Epidémiosurveillance en Santé Végétale (Plateforme ESV), INRAE a décidé d'héberger la majeure partie de l'équipe opérationnelle de la Plateforme ESV dans l'unité BioSP et de confier à BioSP la création de cette équipe (l'équipe OPE). Cette opération, qui s'est matérialisée par le recrutement de 7 ingénieurs permanents et d'une gestionnaire d'unité supplémentaire, a conduit au rattachement de l'unité BioSP au département Santé des Plantes et Environnement (SPE) d'INRAE, en plus du département MathNum qui reste le département pilote de BioSP.

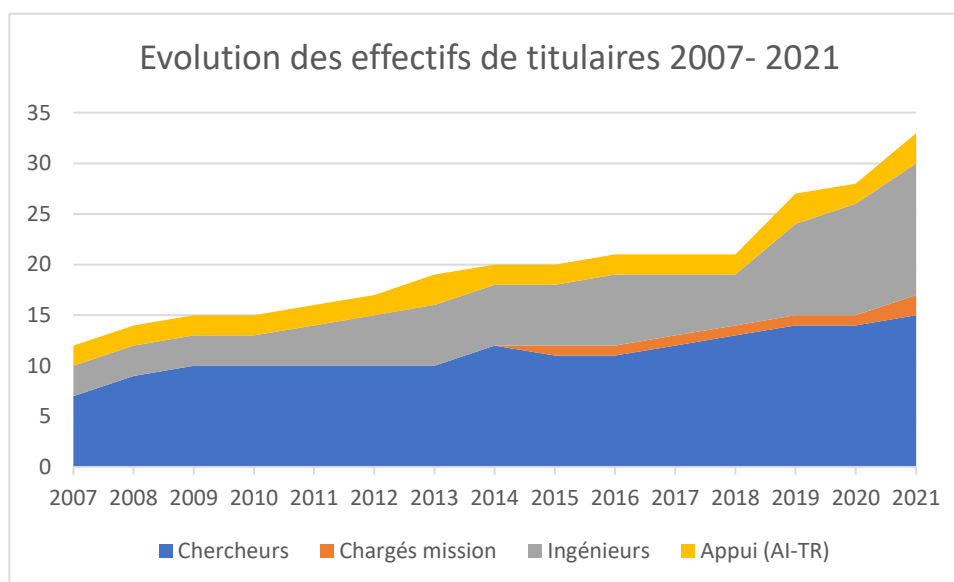


Fig. 1. Évolution des effectifs de l'unité BioSP depuis 2007. La forte inflexion de 2018 coïncide avec la création de l'équipe OPE.

Structuration et thématiques scientifiques de l'unité.

L'unité BioSP développe des travaux en statistique, en systèmes dynamiques, en écologie-épidémiologie, et aux interfaces entre ces différentes disciplines avec un intérêt particulier pour les questions spatiales et spatio-temporelles. Composée depuis sa création et jusqu'en 2018 d'une seule équipe, elle est maintenant constituée de l'équipe Recherche et de l'équipe OPE.

L'unité est structurée scientifiquement autour de 3 axes disciplinaires dans lesquels elle développe des travaux théoriques et appliqués :

1. *Statistique spatiale et spatio-temporelle* : géostatistique ; théorie des champs aléatoires ; géométrie stochastique (processus ponctuels et processus d'objets) ; méthodes pour les données spatio-temporelles ; modèles spatiaux pour les valeurs extrêmes ; méthodes pour les variables du climat.
2. *Systèmes dynamiques déterministes et stochastiques* : analyse des équations aux dérivées partielles (notamment réaction-diffusion) ; équations intégrées-différentielles et autres problèmes non locaux ;

fronts de propagation en milieu hétérogène ; problèmes inverses ; modèles stochastiques non markoviens.

3. *Épidémiologie et écologie spatiales* : distributions d'espèces et communautés ; démographie, cycles de vie et invasions biologiques ; démo-génétique, durabilité des résistances et sauvetage évolutif ; épidémiosurveillance (plantes, abeilles) ; analyse, quantification et cartographie du risque ; échantillonnage basé sur le risque.

Une des forces de l'unité BioSP réside dans sa capacité à faire interagir ces différentes disciplines, pour la mise en place de méthodes innovantes couplant méthodes d'EDP et de statistique, la résolution de questions théoriques en écologie et épidémiologie, la construction de modèles et l'utilisation de données. L'équipe Recherche est impliquée dans chacun de ces 3 axes. L'équipe OPE est davantage impliquée dans les axes 1 et 3.

Depuis 2016, l'unité se structure également autour de priorités scientifiques (PS) sur des enjeux thématiques et méthodologiques, autour desquelles s'organise une animation scientifique, via des groupes de travail, des journées scientifiques et des invitations dans le cadre du séminaire d'unité. Ces priorités sont dynamiques par nature et sont rediscutées annuellement en conseil d'unité. Au 31.12.2021, l'unité compte 5 priorités. Les trois premières ont été définies dès 2016 ; elles correspondent à un important corpus de publications et s'appuient sur des projets (chacune de ces priorités s'appuie notamment sur de gros projets de type ANR et sur un réseau INRAE) : #Epi-stat : Statistique pour l'épidémiologie prédictive ; #Geolearning : Apprentissage statistique pour l'environnement et le climat ; #Evol-sys : Dynamiques évolutives, dynamique de la diversité. Les deux suivantes sont relativement récentes : #Obs-model : Observer et modéliser la distribution d'espèces ; #Spread : Propagations sur supports géométriques complexes. Le **Tableau 1** donne un aperçu de cette structuration croisée de BioSP entre priorités scientifiques et axes disciplinaires.

		Priorités scientifiques				
		#Epi-stat	#Geolearning	#Evol-sys	#Obs-model	#Spread
Axes disciplinaires	Statistique 6.5 ETP	Chercheurs : 1.5 Ingénieurs : 0.5	Chercheurs : 2.5	Chercheurs : 0.5	Chercheurs : 1	Chercheurs : 0.5
	Systèmes dyn. 4 ETP	Chercheurs : 1		Chercheurs : 2		Chercheurs : 0.5 Ingénieurs : 0.5
	Epi./écol. 5.5 ETP	Chercheurs : 1 Ingénieurs : 1		Chercheurs : 1	Chercheurs : 2 Ingénieurs : 0.5	Chercheurs : 0.5

Interactions : **fortes** **ponctuelles**

Tableau 1 : Structuration croisée de BioSP via ses axes disciplinaires et ses priorités scientifiques. Les chiffres indiqués correspondent à une estimation des ETP annuels associés à chaque Axe x Priorité, et ce pour les scientifiques permanents. Cette description n'est pas exhaustive et n'inclut pas les activités transversales (informatique, administration, encadrement, enseignement) et des recherches portant sur des sujets autres, ce qui correspond à 13.5 ETP.

Taille et composition de l'unité au 31/12/2021.

Au 31/12/2021, l'unité compte 32 permanents et 10 non-permanents. La répartition des effectifs entre les différentes catégories de personnels et entre l'équipe recherche et l'équipe OPE est décrite dans le **Tableau 2**. Nous y précisons également le département de rattachement des agents (MathNum, SPE ou ECODIV). L'organigramme de l'unité au 31/12/2021 est présenté en **Fig. 2**.

	Techniciens	Ingénieurs	Chercheurs	Non-permanents
Appui	3 (3 TR MathNum)			
Recherche		6 (3 IR + 2 IE MathNum, 1 IE SPE)	16 (6 CR + 5 DR MathNum, 1 CR + 1 DR SPE, 1DR Ecodiv, 2 CM)	9 (5 doctorants, 1 post- doc, 3 CDD ingénieurs)
OPE		7 (1 IR + 2 IE MathNum, 2 IR+ 2 IE SPE)		1 CDD ingénieur

Tableau 2 : Effectifs au 31/12/2021. CM = chargé de mission.

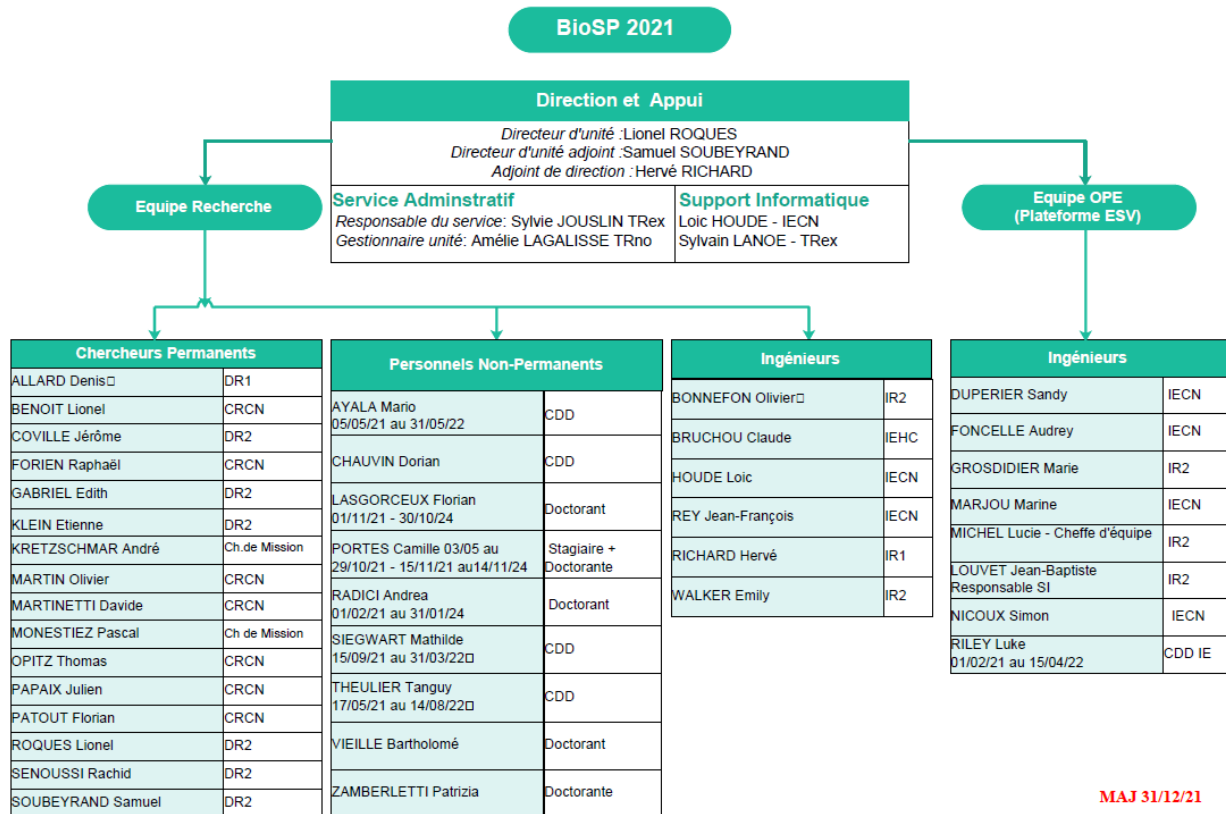


Fig. 2. Organigramme de l'unité au 31/12/2021.

2- PROFIL D'ACTIVITÉS

Activités Equipe Recherche	Répartissez 0 à 100 sur ces 7 items
Administration de la recherche (responsabilité de pilotage de la recherche (VP, Direction d'Institut, DAS, ...) participation à des instances d'évaluation (CNU, CoNRS, CSS...), responsable de volet IdEx, direction de projets -ANR, Horizon Europe, ERC, CPER-, responsabilités éditoriales dans des revues ou collections nationales et internationales)	10
Dissémination de la recherche (partage de connaissances avec le grand public, médiation scientifique, interface science/société)	5
Encadrement de la recherche (implication au niveau D ainsi que dans le suivi de projets post-doctoraux)	10
Contribution à l'adossement d'enseignements innovants à la recherche (EUR, SFRI, etc.)	5
Expertise technique (pouvoirs publics aux niveaux national et régional, entreprises, instances internationales (FAO, OMS, ...))	5
Recherche	60
Valorisation, transfert, innovation	5
Le total doit impérativement être égal à 100	

Activités Equipe OPE	Répartissez 0 à 100 sur ces 7 items
Administration de la recherche (responsabilité de pilotage de la recherche (VP, Direction d'Institut, DAS, ...) participation à des instances d'évaluation (CNU, CoNRS, CSS...), responsable de volet IdEx, direction de projets -ANR, Horizon Europe, ERC, CPER-, responsabilités éditoriales dans des revues ou collections nationales et internationales)	0
Dissémination de la recherche (partage de connaissances avec le grand public, médiation scientifique, interface science/société)	5
Encadrement de la recherche (implication au niveau D ainsi que dans le suivi de projets post-doctoraux)	0
Contribution à l'adossement d'enseignements innovants à la recherche (EUR, SFRI, etc.)	0
Expertise technique (pouvoirs publics aux niveaux national et régional, entreprises, instances internationales (FAO, OMS, ...))	65
Recherche	10
Valorisation, transfert, innovation	20
Le total doit impérativement être égal à 100	

3- ENVIRONNEMENT DE RECHERCHE

BioSP est rattachée aux départements INRAE MathNum (département historique et pilote de BioSP) et SPE.

D'un point de vue territorial, sur le centre INRAE PACA, BioSP contribue principalement à 3 des 4 pôles structurants : le nouveau pôle "M2Sud" animé par BioSP et regroupant les unités ayant une activité en mathématiques/statistique ou en modélisation, ainsi que les pôles "Adaptation aux Changements Globaux" et "Production Horticole Intégrée". L'environnement universitaire de l'unité se partage principalement entre Aix-Marseille Université et Avignon Université.

BioSP participe à l'EUR IMPLANTEUS (Programme interdisciplinaire sur la production et la transformation de plantes méditerranéennes, environnement, santé humaine et durabilité) coordonné par Avignon Université et le centre INRAE PACA (démarrage en septembre 2021). Elle est membre de la FRUMAM (Fédération de Recherche des Unités de Mathématiques de Marseille) depuis 2020. L'unité est également membre de la FR ECCOREV (Fédération de Recherche Ecosystèmes Continentaux et Risques Environnementaux), également centrée sur Aix-Marseille.

BioSP participe pour INRAE à la Plateforme d'Épidémiologie en Santé Végétale (ESV) qui est une structure multi-partenaires incluant le Ministère de l'Agriculture et de l'Alimentation, INRAE, l'Anses, l'ACTA-les instituts techniques agricoles, l'APCA, FREDON France et le Cirad. Cette plateforme a été créée en 2018 par la signature d'une convention cadre pour 10 ans.

4- PRISE EN COMPTE DES RECOMMANDATIONS DU PRÉCÉDENT RAPPORT

Avis global

Recommandation : *Le choix fait de conserver une structure globale non différenciée en équipe s'avère tout à fait pertinent et productif. Il semble clair au comité d'experts que cette circulation interne serait encore bien améliorée par le recrutement d'un chercheur en probabilités et processus stochastiques pouvant créer une nouvelle interface entre systèmes dynamiques et statistiques spatio-temporelles. Ce recrutement était du reste déjà affiché comme une priorité dans le précédent rapport et on ne peut que souhaiter qu'il soit effectif à très brève échéance.*

Réponse : L'arrivée de l'équipe OPE a conduit l'unité à se structurer en deux équipes. Néanmoins, en cohérence avec le précédent rapport d'évaluation, l'unité a décidé collectivement de conserver une unique équipe "Recherche". L'unité a poursuivi sa demande de recrutement d'un chercheur en probabilités et processus stochastiques tout en renforçant ses compétences dans ce domaine. Cela lui a permis d'effectuer le recrutement en 2018 de Raphaël Forien qui a suivi un cursus "Math-Bio" à l'ENS Paris et mené une thèse en probabilités, en partie effectuée à l'université d'Oxford. Le profil sur lequel il a été sélectionné visait à recruter un chercheur apportant des compétences théoriques sur les modèles aléatoires spatio-temporels (modèles particuliers, processus de branchement, EDS, EDPS), venant compléter la panoplie d'outils méthodologiques de l'unité BioSP en se situant à l'articulation des deux grands axes de recherche actuels que sont les systèmes dynamiques et les statistiques spatiales. Nous remercions le précédent comité d'évaluation HCERES pour son soutien sur ce point.

Recommandation : *(augmenter le) nombre d'interactions directes avec l'environnement non académique, où il y a pourtant des demandes évidentes qui pourraient en particulier donner lieu à des financements de thèses.*

Réponse : Nos partenariats non académiques sont souvent soit indirects via d'autres unités INRAE, soit tournés vers des structures semi-publiques ou publiques (e.g. ANSES, Ministère de l'Agriculture et de l'Alimentation (MAA), filière apicole...) ; la Plateforme ESV est à cet égard un élément structurant majeur. Depuis 2020, Denis Allard porte le projet de chaire de mécénat "Geolearning", avec l'Ecole des Mines de Paris. Cette chaire, dont les travaux démarreront en 2022, sera financée par quatre entreprises mécènes intéressés par des problématiques en lien avec les données spatio-temporelles, les événements extrêmes et le Machine Learning. Voir le Dom 4 pour plus d'éléments.

Critère 1 : qualité et production scientifiques

Recommandation : *(remédier à) la fragilité numérique de la thématique « Systèmes Dynamiques et EDP ».*

Réponse : Dès 2017, à l'issue du précédent rapport, l'unité a porté un projet de recrutement de CRCN sur le thème "Equations aux dérivées partielles : théorie, modèles et aspects numériques." Ce profil a été arbitré positivement par les tutelles et a conduit au recrutement en 2021 de Florian Patout, qui avait fait une thèse à l'ENS Lyon puis un post-doc de 2 ans dans l'unité BioSP avant son recrutement. Notons également le passage DR2 de Jérôme Coville et les deux années de délégation que Matthieu Alfaro (PR, Rouen) a passées à BioSP. Ces éléments, ainsi que le recrutement du probabiliste Raphaël Forien qui participe activement à cette thématique, garantissent une pérennité à la thématique "*Systèmes Dynamiques et EDP*" et démontrent la confiance accordée par les tutelles concernant le développement de cette thématique sur laquelle BioSP devient un acteur important à INRAE.

Recommandation : *(remédier à) une pyramide des âges déséquilibrée.*

Réponse : Les nombreux recrutements de scientifiques effectués au cours de la période dans l'équipe Recherche (Davide Martinetti, CRCN SPE, Raphaël Forien CRCN MathNum, Florian Patout CRCN MathNum, Lionel Benoit CRCN MathNum, Emily Walker IR2, Edith Gabriel DR2) et la création de l'équipe OPE ont permis de nettement rééquilibrer la pyramide des âges de BioSP. L'âge moyen des permanents est de 41 ans en 2021 contre 47 ans en 2015.

Recommandation : *Les chercheurs de BioSP sont incités à diversifier les journaux dans lesquels ils publient leurs résultats de méthodologie statistique.*

Réponse : En raison du positionnement de la composante statistique de l'unité sur les statistiques spatiales et les événements extrêmes, les principales revues cibles sont (à juste titre nous semble-t-il) des journaux comme *Spatial Statistics*, *Extremes*, *Stochastic Environmental Research and Risk Assessment* et *Environmetrics*. Néanmoins, sur la période considérée, on note également plusieurs publications dans des journaux statistiques généralistes à très forte notoriété comme *Annals of Applied Statistics*, *Bernoulli*, *Electronic Journal of Statistics*, *JASA*, *Scandinavian Journal of Statistics* et *Statistics and Computing*.

Recommandation : *Une réflexion reste à mener à présent sur le développement logiciel : quel devenir pour les codes, et quelle utilisation hors de l'unité ? Quelle part donner aux outils ponctuels ou génériques ?*

Réponse : L'unité a tâché d'instaurer des bonnes pratiques en matière de développement logiciel : notamment l'utilisation d'outils de versionnement (GitLab auto-hébergé, GitHub) pour permettre le travail collaboratif, le partage et l'archivage de code ponctuel, générique ou spécifique ; l'utilisation de répertoires ouverts de logiciels et données (Zenodo, Open Science Framework, dépôts GitLab publics) quand ceux-ci sont adossés à des publications. Plusieurs projets informatiques / logiciels ayant vocation à rayonner en dehors de l'unité ont été développés pendant la période : le projet SK8 visant une infrastructure matérielle et logicielle de développement et d'hébergement d'applications Web R-Shiny, le projet MSE autour de la mise en place d'un environnement informatique dédié à l'approche mécanistico-statistique (cf Dom 3 Ref 1 équipe Recherche) ou encore le package R et l'interface web "landsepi" [LOG-RRP21] permettant la simulation d'épidémies à l'échelle d'un territoire agricole. L'unité encourage des porteurs de projets informatiques nécessitant un travail d'ingénierie conséquent et pérenne à l'appuyer sur les CATI (Centre Automatisé de Traitement de l'Information) qui regroupent des forces d'ingénieurs au niveau de l'institut. C'est le cas pour les projets MSE et SK8.

Critère 2 : rayonnement et attractivité académiques

Recommandation : *(remédier au) peu d'invitations de collaborateurs pour des séjours longs ou semi-longs.*

Réponse : L'unité a invité trois chercheurs sur des périodes longues via des accueils en délégation : Edith Gabriel (statisticienne, 2014-2017), Florent Bonneau (statisticien, 2019-2021) et Matthieu Alfaro (EDPiste, 2019-2021). Symétriquement, Denis Allard a été invité comme professeur associé en 2016 puis en 2019 (deux mois à chaque fois) à l'Université Ca'Foscari de Venise.

Recommandation : *Des membres de l'unité ont la capacité de porter des projets européens de type ERC et (...) les encourage à mener une réflexion en ce sens.*

Réponse : L'unité porte un projet de chaire de mécénat (co-porté par l'équipe "géostatistique" de l'Ecole des Mines de Paris) ayant pour intitulé: "Geolearning: géostatistique, événement extrême et Machine Learning

pour la transition écologique”. Cette chaire d’une durée de 5 ans sera financée à hauteur de 1.8 M€ par 4 partenaires mécènes: ANDRA, BNP-paribas, CCR et SCOR.

Recommandation : *Le site internet de l’unité pourrait être enrichi et mieux illustré, et avoir une version en langue anglaise.*

Réponse : Le site internet de l’unité a été entièrement remodelé avec notamment l’ajout de contenus dynamiques, la présentation des 2 équipes, un trombinoscope, un agenda du séminaire de l’unité, la présentation de publications récentes. Afin de garantir une bonne cohérence entre les versions en anglais et en français du site, nous avons fait le choix d’une traduction automatique, qui donne actuellement des résultats satisfaisants.

Critère 3 : interactions avec l’environnement économique, social, culturel et sanitaire

Recommandation : *L’unité devrait se saisir de la problématique des relations directes avec des partenaires industriels, notamment locaux. La production logicielle est très bonne et peut contribuer à améliorer cette visibilité.*

Réponse : cf réponse à l’avis global

Critère 4 : organisation et vie de l’unité

Recommandation : *Le comité d’experts a jugé qu’un axe « ingénierie » transversal aurait toute sa place au sein de l’unité BioSP, pour permettre une meilleure identification des compétences et savoir-faire, et ainsi peut-être une répartition plus efficace des tâches.*

Réponse : L’unité a longtemps fonctionné avec un axe « Ingénierie pour l’analyse de modèles numériques » et plusieurs évaluations passées regrettaient une connexion trop faible entre recherche et ingénierie, contrairement à la recommandation du rapport de 2017. Le projet d’unité cherche donc volontairement à distribuer l’ingénierie équitablement dans les différents axes disciplinaires et priorités scientifiques, afin d’éviter que l’ingénierie ne s’enferme dans un axe déconnecté des activités de recherche de l’unité. Ceci dit, la création de l’équipe OPE en 2018, sur des missions d’ingénierie opérationnelles clairement distinctes des missions de recherche du reste de l’unité, a de fait généré un regroupement partiel des ingénieurs dans une structure infra-unité. De plus, en cohérence avec les conseils du comité HCERES, nous avons renforcé l’importance et l’efficacité des structures d’animation concernant les domaines du calcul, de l’informatique et des données avec la mise en place de six groupes de travail sur ces thèmes, animés pour la plupart par des ingénieurs de l’unité. Enfin, les ingénieurs de BioSP (des équipes Recherche et OPE) sont impliqués dans la structure transversale qu’est le CATI IMOTEP, où ils peuvent échanger avec les ingénieurs d’autres unités et mener des projets fédérateurs au sein de l’unité et au-delà, comme le projet SK8 visant une infrastructure matérielle et logicielle de développement et d’hébergement d’applications Web R-Shiny.

Critère 5 : implication dans la formation par la recherche

Recommandation : *(remédier à) un nombre de doctorants toujours faible dans un contexte structurel difficile ; pas de suivi systématique du devenir des docteurs de l’unité. L’unité pourrait s’investir dans des « écoles jeunes chercheurs » ; un investissement en formation accru dans les écoles doctorales (SAS et MIM) sur lesquelles l’unité a décidé de se concentrer pourrait apporter de la visibilité.*

Réponse : Lors de la précédente évaluation, les doctorants émergeaient à 5 écoles doctorales différentes. Depuis 2016 nos doctorants relèvent essentiellement de 2 écoles doctorales (l’ED 184 Mathématique et Informatique d’Aix-Marseille Université et l’ED 536 Agrosociences / Sciences Agronomiques d’Avignon Université). L’unité BioSP est devenue équipe d’accueil de l’ED 184 MIM en janvier 2016, puis membre de la FRUMAM (Fédération de Recherche des Unités de Mathématiques de Marseille) en janvier 2020. Sur la période correspondant à la précédente évaluation, l’unité avait accueilli 13 doctorants, vs 17 pour la période actuelle, auxquels il faut ajouter les doctorants co-encadrés par des chercheurs de l’unité mais hébergés dans un autre laboratoire (8 doctorants, comptabilisés dans l’Annexe 1). En ne comptabilisant que les étudiants accueillis à BioSP, on arrive à un ratio doctorant/chercheur de 0.4 sur la période 2016-2021, supérieur à la moyenne INRAE, qui est d’environ ¼ d’après le bilan social annuel. La possibilité d’attirer les meilleurs étudiants reste un sujet important pour BioSP, que nous discutons ci-dessous (cf. synthèse Dom 2).

DOCUMENT D'AUTOÉVALUATION

Dans la suite de ce document, les éléments entre crochets, p ex [PUB-CAM21] ou [LOG-BRL21] font référence au tableau « Données de production et d'activités ». Nous faisons référence aux différentes sections du document en indiquant "Dom X Ref Y" pour diriger vers la Référence Y du Domaine d'évaluation X.

Domaine d'évaluation 1 : Profil, ressources et organisation de l'unité

Référence 1. L'unité possède des ressources adaptées à son profil d'activités et à son environnement de recherche.

Conformément à leurs missions, l'équipe Recherche présente un profil d'activité essentiellement tourné vers la recherche alors que l'équipe OPE a un profil d'activité tourné vers l'expertise et le transfert.

Les financements récurrents de l'unité s'élèvent en moyenne à 142k€/an sur la période, auxquels sont venus s'ajouter 143k€ en 2019 pour l'installation de l'équipe OPE. Une part non négligeable de ces financements récurrents est obtenue via des appels d'offres internes à INRAE (dont 29k€/an de soutien des départements aux réseaux/CATI et 15k€/an sous forme d'appels à projets compétitifs auprès des départements). Une partie importante de la dotation récurrente (celle qui ne correspond pas aux appels d'offres internes) est mutualisée. Cela permet de financer l'achat de matériels et logiciels collectifs (notamment le financement d'une partie du cluster de calcul, ou l'achat de jetons pour des logiciels) et individuels, l'invitation de scientifiques au séminaire hebdomadaire (café des sciences, <https://biosp.mathnum.inrae.fr/cafe-sciences>), l'invitation de scientifiques sur des périodes plus longues, le financement de missions, de stages et de CDD. Depuis 2019, nous avons mis en place une procédure de financement de projets internes à l'unité. Cela a notamment permis de financer un CDD de 16 mois sur un projet d'ingénierie (MSE, cf Dom 3 Ref 1 de l'équipe Recherche).

Les appels d'offres internes INRAE sont des outils d'animation importants pour l'unité. Ainsi, BioSP a porté sur la période (ou co-animé) 4 « réseaux d'animation scientifique », chacun des 4 réseaux ayant été renouvelé plusieurs fois sur une base annuelle ou bi-annuelle (cf Dom 2 Ref 1 pour des détails supplémentaires concernant ces réseaux).

Au-delà de cette dotation récurrente, l'unité dispose d'importantes ressources propres, via des contrats européens et ANR notamment. Au total, ces financements représentent 306k€/an sur la période. Parmi ces projets, ceux portés par BioSP intègrent souvent un nombre important de collègues de l'unité. C'est par exemple le cas du projet BEYOND (5 chercheurs et 5 ingénieurs) ou du projet RESISTE (5 chercheurs et 2 ingénieurs).

Avec l'appui du Ministère de l'Agriculture et de l'Alimentation (MAA) au moment de la création de l'équipe OPE (coïncidant avec la création de la Plateforme ESV), l'unité a fait construire une extension permettant d'héberger 10 agents. Cela permet aux permanents de ne pas être plus de 2 par bureau, sans avoir de contrainte de place pour l'accueil des non-permanents. Cette extension n'est pas une structure pérenne, et à terme (2023 ou 2024), un déménagement est prévu dans un bâtiment du centre présentant une superficie bien supérieure à celle de nos locaux actuels.

Concernant les aspects matériels et logiciels, les ressources de l'unité permettent aux agents de renouveler leur poste de travail portable tous les 3 ans s'ils le souhaitent et d'acquérir toute licence logicielle nécessaire à leur travail (actuellement, nous disposons par exemple de licences Matlab, pycharm, Adobe Creative Suite, Office, licence Only Office pour la co-édition de documents sur le cloud de la Plateforme ESV et de l'unité, API Google Translate et ScaleSERP pour la Plateforme ESV). L'unité a également souscrit à un abonnement collectif sur Overleaf. Les agents de l'unité ont accès au cluster de calcul de l'unité (cf. Dom 2 Ref 4) et bénéficient de l'appui des ingénieurs informaticiens de l'unité pour son utilisation et sa montée en gamme en termes de capacité et de performance.

Concernant les ressources documentaires, INRAE fournit un accès aux principales revues internationales. Nous disposons d'abonnements complémentaires à ceux d'INRAE, notamment à des revues scientifiques et informatiques grand public (La Recherche, Sciences et Avenir, Linux, Programmez). L'unité dispose

également d'une bibliothèque comportant plusieurs centaines d'ouvrages (950). Cette collection est régulièrement mise à jour sur proposition des chercheurs et ingénieurs de l'unité.

Référence 2. L'unité s'est assigné des objectifs scientifiques, y compris dans la dimension prospective de sa politique.

De par ses activités en statistique et en mathématique, l'unité contribue essentiellement à deux des trois grands objectifs scientifiques du département MathNum (schéma stratégique 2021-2025) : GOS 1 "Maîtriser les méthodes pour acquérir, gérer et intégrer données et connaissances face à la multiplication des sources d'information" et GOS 2 "Développer les méthodes de modélisation et d'analyse en vue de comprendre et anticiper les trajectoires de systèmes complexes" avec des intérêts thématiques dans 2 des 5 champs thématiques du département : CT3 "Probabilités, statistique et apprentissage automatique" et CT4 : "Modélisation des systèmes complexes et systèmes dynamiques". Ses recherches en écologie et en épidémiologie contribuent quant à elles aux GOS et aux CT du département SPE, notamment le GOS 1.2 "Réduire la vulnérabilité sanitaire des paysages agricoles : comprendre l'adaptation des populations de bioagresseurs ; développer l'épidémiologie ; imaginer des stratégies durables pour cultiver des peuplements résistants" et le GOS 4.1 "Comprendre et anticiper les invasions biologiques et les crises sanitaires : développer une systématique intégrative ; caractériser la capacité invasive des organismes et les conditions d'émergence dans un nouvel environnement ; modéliser pour anticiper le risque et optimiser les stratégies de gestion ; développer la gestion prophylactique pour vivre avec les organismes envahissants" via à la fois l'équipe Recherche et l'équipe OPE.

Au niveau de l'Institut, INRAE organise ses recherches selon 5 grandes orientations scientifiques (OS), qui articulent des priorités de recherche portant des dimensions académiques, d'innovation et de politique publique. De par ses activités, l'unité BioSP contribue à 4 de ces OS.

- Notamment l'OS 1 "Répondre aux enjeux environnementaux et gérer les risques associés" via les priorités scientifiques #Geolearning et #Evol-sys qui font respectivement écho aux OS 1.1 et 1.4 "*Changement climatique : intégrer les démarches d'atténuation et d'adaptation*" / "*Evaluation et gestion des risques naturels et climatiques*" et à l'OS 1.3 "*Compréhension et mobilisation des mécanismes d'adaptation du vivant pour la sélection génétique et la préservation de la biodiversité*".
- Également l'OS 2 "Accélérer les transitions vers des systèmes agricoles et alimentaires agroécologiques en tenant compte des enjeux économiques et sociaux", avec des travaux en lien direct avec l'OS 2.2 "*Progression vers des agricultures sans pesticides de synthèse*" réalisés par exemple dans le cadre de projets inscrits dans les PIA et France Relance comme BEYOND et SEPIM (en lien étroit avec la PS #Epi-stat) et illustrés par un certain nombre d'articles étudiant l'utilisation de la diversité génétique ou le biocontrôle pour réduire les pesticides et de nombreux chapitres d'ouvrages de synthèse et de vulgarisation sur le sujet.
- L'OS 4 "Favoriser une approche globale de la santé", essentiellement autour du point 4.1 *Emergences et re-émergences des maladies transmissibles, au sein et entre les systèmes environnementaux, agricoles et alimentaires* via les activités de l'équipe OPE, la PS #Epi-stat, et le réseau méthodologique ModStatSAP (cf Dom 2 Ref 1).
- L'OS 5 "Mobiliser la science des données et les technologies du numérique au service des transitions" et notamment l'OS 5.1 "*Systèmes complexes et évolutifs*" qui par définition couvre une partie importante de nos activités (notamment l'Axe Systèmes Dynamiques et la PS #Evol-Sys), et l'OS 5.2 "*Capteurs et systèmes d'acquisition d'information*" qui entre en résonance avec le projet sur lequel L Benoit a été recruté en 2021 comme CRCN et les PS #Geolearning et #Obs-Model.

L'unité BioSP est impliquée dans de nombreuses collaborations avec les autres unités du centre INRAE PACA (en particulier, les unités Ecologie des Forêts Méditerranéennes, Pathologie Végétale, Plantes et Systèmes de culture Horticoles, Abeille et Environnement), ce qui structure en partie ses activités de recherche les plus appliquées. D'un point de vue académique, ses interactions principales sont :

- Aix-Marseille Université : interactions importantes avec les mathématiciens de l'Institut de Mathématiques de Marseille (I2M), notamment les équipes "Analyse Appliquée" et "Mathématiques de l'Aléatoire" avec 16 publications co-signées sur la période évaluée, le co-encadrement de 4 thèses, l'inscription de 9 doctorants à l'ED Mathématiques et Informatique de Marseille (MIM, thèses orientées mathématiques), l'inscription de 4 chercheurs HDR de BioSP à cette ED, des enseignements dispensés par les chercheurs de BioSP dans des masters de l'I2M. Les domaines de recherche concernés sont essentiellement l'analyse des EDP (elliptiques, paraboliques, intégro-différentielles) à travers des collaborations "historiques" (depuis 2005) avec l'équipe d'Analyse Appliquée de l'I2M, les probabilités (collaborations avec E Pardoux depuis le recrutement de Raphaël Forien) et la statistique (collaboration avec D Pommeret).
- Avignon Université : via de nombreuses interactions avec les statisticiens du Laboratoire de Mathématiques d'Avignon (LMA) avec 6 publications co-signées sur la période évaluée, l'accueil d'enseignant chercheur en délégation (2 accueils de MCF sur la période évaluée), le co-encadrement d'une thèse, l'inscription de 6 doctorants à l'ED Sciences et Agrosociétés (SAS, thèses orientées biologie pour la plupart mais pas exclusivement), l'inscription de 3 chercheurs HDR de BioSP à cette ED. La proximité géographique du LMA, l'accueil en délégation d'Edith Gabriel et de Florent Bonneau puis le recrutement d'Edith, le partage d'un agenda concernant les séminaires d'unité sont autant d'éléments qui ont favorisé des collaborations très régulières. Ces collaborations concernent essentiellement les statistiques spatiales et spatio-temporelles.
- MINES Paris : l'unité entretient une collaboration de longue durée avec l'équipe "Geostatistique" du Centre de Géosciences de MINES Paris qui se traduit par des enseignements (option Géostatistique et Probabilité Appliquée), des co-encadrements de thèses (une thèse soutenue, une thèse en cours, cf Annexe 1), plusieurs publications communes et par le montage en commun de la Chaire de mécénat Geolearning. Cette chaire, dont le montage a débuté en février 2020, démarrera en 2022 pour durer jusqu'en 2027. Par son financement (1.8 M€ sur 5 ans) et par ses thèmes, cette chaire jouera un rôle très structurant pour les liens entre BioSP et MINES Paris dans les années à venir.

Membre de la FRUMAM depuis 2020 (Fédération de Recherche des Unités de Mathématiques de Marseille), à laquelle émerge également le LMA et l'I2M, BioSP organise à ce titre plusieurs rencontres scientifiques chaque année dans les locaux de la FRUMAM (Marseille, site St-Charles).

L'ensemble des personnels de l'unité se réunit mensuellement pour une "réunion d'unité" (qui vaut conseil d'unité, conseil scientifique et organe représentatif des personnels), dont l'ordre du jour est variable et dépend du contexte. La plupart des décisions en matière de RH et d'orientation scientifique sont prises de façon collégiale au cours de ces réunions. Parmi les questions ayant un fort effet structurant, il s'agit notamment de préparer les futurs recrutements en identifiant collectivement les besoins, en recueillant les propositions des différents agents et en les hiérarchisant. Jusqu'à présent, ces décisions ont toujours pu se prendre de façon collective, sans engendrer de mécontentement manifeste. Typiquement, nous faisons remonter 2 demandes annuelles de postes, et le classement retenu avantage le plus souvent la demande la plus ancienne. Nous suivons le même processus pour interclasser les demandes de bourses de thèses qui sont faites aux départements. Ces réunions permettent également de définir l'organisation collective de l'unité, notamment sa politique en matière de télétravail (dans le cadre des possibilités offertes par INRAE), l'animation hebdomadaire, l'accueil des nouveaux, la répartition des personnels dans les locaux.

Parallèlement à ces rencontres, nous organisons une AG annuelle hors-les-murs qui est à la fois un moment de convivialité et une opportunité de réfléchir à notre politique scientifique à moyen et long terme. C'est notamment au cours de ces rencontres que sont (re)définies les priorités scientifiques de BioSP (**Tableau 1** et **Fig. 3** du Dom 3 Ref 1 ci-dessous).

Depuis 2021, nous avons mis en place six groupes de travail sur des sujets stratégiques pour l'organisation de l'unité dans les domaines du calcul, de l'informatique et des données : (1) Le groupe "Cluster de calcul" (qui pré-existait avant 2021) sur l'infrastructure matérielle et logicielle de calcul de BioSP ainsi que le suivi des besoins et l'accompagnement des utilisateurs ; (2) Le groupe "Outils et méthodes numériques pour l'analyse mécanistico-statistique" ; (3) Le groupe "Systèmes d'information" (SI) sur la montée en gamme et la mise en cohérence des SI BioSP et Plateforme ESV ; (4) Le groupe "AppliWeb" sur les modalités de déploiement des applications informatiques "Web" de BioSP ; (5) Le groupe "Postes de travail" sur la gestion des postes

informatiques individuels ; (6) Le groupe "Données scientifiques" sur le partage des pratiques relatives à l'utilisation des données génériques au sein de BioSP. Ces groupes sont animés pour la plupart par les ingénieurs de l'unité. Il sont un lieu d'échange privilégié entre eux et les utilisateurs des ressources informatiques.

Référence 3. Le fonctionnement de l'unité est conforme aux réglementations en matière de gestion des ressources humaines, de sécurité, d'environnement et de protection du patrimoine scientifique.

L'unité est fortement engagée dans l'accompagnement des parcours professionnels individuels et des collectifs de travail. Chaque année, chaque agent bénéficie des dispositifs proposés par l'institut (formations individuelles, collectives, diplômantes). A partir des besoins exprimés de chacun, titulaires et non titulaires, l'unité étudie et élabore, accompagné en cela par un référent formation au sein de l'unité (à savoir Jérôme Coville sur la période évaluée), un plan de formation annuel favorisant le développement des compétences et l'évolution des carrières des agents. La plupart des années, l'unité finance quelques formations en plus de celles prises en charge par la "formation permanente" du centre PACA, et ce en fonction des demandes des personnels.

Dans le cadre du développement de programmes structuraux, de la construction de l'espace européen de la recherche ou de l'évolution des techniques, les formes d'accompagnement se sont diversifiées (écoles chercheurs, accompagnement de la mise en réseau des compétences et des communautés de recherche). Des ressources pédagogiques multimédia et dispositifs de e.formation sont également à disposition dans différents domaines : sciences, informatique, prévention, management, ...

En matière de conditions de travail, l'unité veille à ce que chaque agent travaille dans des locaux entretenus, dispose d'un espace de travail confortable, d'un mobilier ergonomique et d'un matériel informatique performant. Les droits à congé sont respectés, le télétravail est favorablement adopté. La moindre absence prolongée et ou inhabituelle d'un agent est très vite détectée et signalée. Depuis 2021, le travail en présentiel 3 jours sur 5 est fortement encouragé (politique interne à l'unité, cf fin de ce paragraphe) afin de garantir les interactions et les échanges professionnels. Les agents de l'unité sont invités à choisir les jours de télétravail en cohérence avec leurs principaux collaborateurs dans l'unité. Par ailleurs le lundi, jour des réunions d'unité mensuelles et des séminaires, n'est pas télétravaillable.

Comme dans toutes les unités du centre INRAE PACA, les agents disposent des services de l'assistante sociale, d'un suivi médical auprès du médecin de prévention, des compétences du service prévention centre puis de l'appui de l'agent de prévention (Olivier Bonnefon sur la période évaluée) missionné par le DU et chargé de relayer tout problème en matière de sécurité individuelle ou collective.

En matière de prévention des risques psycho-sociaux, l'unité porte naturellement son attention sur les risques liés au travail à distance (isolement, mobilier de travail inadapté, charge de travail non contrôlable..) mais aussi sur le partage des valeurs, l'intégration sociale et professionnelle et les formes de discrimination. Courant 2021, le Centre PACA s'est engagé dans un projet Diversité-Egalité, mettant en place des journées formation pour les DU, DUA et les personnes chargées de les représenter avec pour objectifs : connaître la réglementation, partager des situations managériales, travailler sur les biais, échanger sur la réalité du vécu en tant que manager et acteurs RH et définir ensemble les priorités d'actions à développer pour le centre PACA. Deux agents de l'unité (Lucie Michel et Sylvie Jouslin) ont suivi cette formation et ont organisé une discussion au sein de l'unité en juin 2021.

En matière informatique, conformément aux préconisations de la DSI INRAE, l'unité a localisé une partie de son SI sur le Datacenter INRAE de Toulouse pour garantir un environnement adapté et sécurisé selon l'état de l'art : l'ensemble des services et applications logicielles sont déployés sur des machines virtuelles mises à disposition par la DSI et pour lesquelles celle-ci assure l'infrastructure matérielle et la sauvegarde.

La Plateforme ESV, en coopération avec la RSSI INRAE, est entrée dans une démarche de cybersécurité visant à terme à faire homologuer son SI. Une cartographie exhaustive du SI ainsi qu'une analyse de risque ont été réalisées, permettant d'évaluer le risque et de mettre en place un plan d'action actuellement en cours de déploiement.

La partie du SI locale à l'unité est hébergée sur des serveurs de virtualisation de type PROXMOX. Leur architecture est basée sur une duplication logicielle et/ou matérielle des différents services et applications. Celles-ci sont protégées par un proxy et celles ayant le plus de criticité disposent d'un plan de reprise.

Afin de garantir la pérennité, la traçabilité et le partage, les fonds documentaires de l'unité et de la Plateforme ESV sont déployés sur un serveur NextCloud via un prestataire respectueux de la confidentialité. Ce service est hébergé sur un ensemble de VM provisionnées dans un datacenter OVH localisé en France.

La mise en conformité des postes de travail (PdT), avec en particulier leur chiffrement pour les postes de la Plateforme ESV, est réalisée par le technicien PRI (Personnes Ressources Informatique) dernièrement recruté. L'ensemble des postes de travail possède un système de sauvegarde automatique vers un volume S3 localisé en datacenter INRAE. Enfin, l'unité s'est organisée en groupe de travail par thématiques (SI, Données, PdT, AppliWeb...) pour garantir que sa stratégie globale est adaptée à ses besoins et rester en cohérence avec le plan données pour la science en cours de déploiement à INRAE, cf. Dom 1 Ref 2. Concernant ce plan et sa mise en oeuvre, deux "référents données opérationnels" ont été nommés pour l'unité et servent de courroie de transmission avec la DipSO (direction pour la science ouverte ; <https://www6.inrae.fr/dipso/>) et avec les "référents données stratégiques" nommés par les départements.

L'unité s'est engagée dans une démarche éco-responsable pendant la période évaluée en s'impliquant dans diverses actions :

- gestion des déchets avec le tri sélectif pour les emballages dans l'espace cafétéria et l'installation d'un composteur à proximité du bâtiment pour les déchets organiques. Le tri des papiers, cartons, cartouches d'encre, D3E, stylos, piles, ampoules et tube fluos est réalisé avec sérieux dans l'unité, organisé par les services d'appui à l'échelle du centre (SDAR).
- gestion des consommations énergétiques avec une mission d'audit énergétique du bâtiment commandée et payée par BioSP auprès du bureau d'étude INGEFLUX. Cette étude a permis en particulier de décomposer la consommation électrique du bâtiment entre les différents postes principaux (salle machine, chauffage et climatisation, autres) et de consolider les suivis mensuels de consommation transmis par les SDAR.
- connaissance de notre empreinte carbone avec un premier bilan CO2 en grandes masses dès 2019 et un bilan plus rigoureux sur l'outil GES1point5 pour 2021 (appuyé sur les chiffres consolidés de l'audit). La direction a fait le choix de reproduire ce bilan GES1points5 annuellement pour un suivi longitudinal.

La connaissance de l'empreinte carbone permet d'inciter à des comportements plus vertueux sur les postes les plus impactants : mobilité (missions, déplacements domicile-travail, télétravail) (49% de l'empreinte), achats (durée de vie des équipements informatiques) (29% de l'empreinte).

Plusieurs agents sont investis à l'échelle du centre pour la promotion des déplacements à vélo.

Face à des situations d'urgences exceptionnelles, l'unité s'appuie sur les directives du service Prévention et de la Présidence de centre. La création en mars 2020 d'une cellule de crise sur le centre PACA a permis de réorganiser le travail en unité dans les meilleures conditions et dans le respect des consignes sanitaires en période de pandémie Covid-19 et de confinement (télétravail obligatoire ou encouragé suivant les périodes). La mise à jour régulière du guide de prévention dans lequel le PCA (Plan de Continuité d'Activité) est détaillé permet à l'unité et au collectif d'être régulièrement informé des mesures de prévention en vigueur. Pendant les périodes les plus critiques (confinement avec télétravail obligatoire), conformément au PCA, deux agents de BioSP se sont portés "Personnes Volontaires Mobilisables". Ces deux agents pouvaient se rendre sur site, uniquement lorsque cela était nécessaire, afin de pallier aux incidents informatiques nécessitant d'accéder aux machines ou d'accéder aux documents du secrétariat. De fait, la plupart des activités de l'unité sont télétravaillables, et notre principale préoccupation était d'éviter l'isolement des nouveaux arrivants et le délitement des liens entre agents du labo. Pour cela, nous avons organisé des réunions régulières en visio et encouragé les permanents à renforcer leur suivi des stagiaires et des contractuels. L'unité a maintenu une animation scientifique régulière notamment via ses travaux sur l'épidémie de Covid-19 qui ont mobilisé de nombreux agents, cf Dom 3 Ref 1 et Portfolio. Suite à cet épisode, l'unité a décidé collectivement de définir sa propre stratégie concernant le télétravail afin de faciliter une bonne reprise des interactions entre agents. Depuis 2020, INRAE prévoit en effet l'attribution d'un volume de jours de télétravail flottants pouvant aller jusqu'à 100 jours par an (porté à 144 jours en 2022). Au sein de BioSP, nous avons décidé que le lundi ne serait pas télétravaillable, et que les agents devraient définir des jours *fixes* de télétravail (au plus 2 par semaine, plus 10 jours flottants par an). Ces jours fixes sont définis par les agents, et permettent notamment de maintenir

3 jours de présence simultanée entre les non-permanents et leurs encadrants ou encore la présence d'une GU chaque jour de la semaine.

Synthèse de l'autoévaluation

Forces. La préservation d'une qualité de vie au travail optimale, caractérisée par un épanouissement personnel et professionnel de chacun des agents ainsi que des interactions fréquentes entre les agents, est au centre des préoccupations de l'unité. Le modèle de prise de décision collégiale en matière de stratégie scientifique et managériale ainsi que la mutualisation d'une partie importante des ressources participent à cette dynamique positive. La présence de deux gestionnaires d'unité et d'un personnel compétent et en effectif suffisant pour la gestion du SI permettent une utilisation fluide des ressources. Sur le plan financier, l'absence d'expérience ou de matériel très coûteux (en comparaison d'autres unités INRAE) et les succès réguliers aux appels à projets permettent aux agents de bénéficier de ressources suffisantes pour leurs achats de matériel, missions, frais de publication et encadrements de stagiaires. Les thématiques couvertes par l'unité résonnent fortement avec les orientations scientifiques des tutelles, ce qui s'est traduit notamment par des recrutements réguliers (**Fig. 1**) bien répartis entre les axes disciplinaires de BioSP et par l'hébergement de l'équipe OPE de la Plateforme ESV. BioSP est bien implantée localement et interagit fortement avec les unités INRAE du centre ainsi qu'avec les universités AMU et AU.

Faiblesses. Avec l'arrivée de l'équipe OPE, impulsée par les tutelles, l'unité espérait bénéficier de locaux neufs ou d'une extension pérenne de son bâtiment. La politique de gestion du patrimoine immobilier d'INRAE ainsi que les ressources allouées n'ont pas permis ces acquisitions, et nous avons dû investir dans des bâtiments modulaires ("Algecos") pour disposer d'un nombre de postes de travail suffisant. Un déménagement est prévu sur un bâtiment du centre présentant une superficie bien supérieure à celle de nos locaux actuels (environ x2) et qui devrait se libérer à l'issue des travaux sur le centre (devant initialement se terminer en 2021). Néanmoins, l'architecture de ce bâtiment est très particulière et un investissement financier important sera nécessaire pour qu'il permette d'héberger confortablement les agents de BioSP. Sur le plan informatique, la politique des tutelles au début de la période évaluée était de migrer les équipements vers des data-centers. L'unité a choisi de conserver une partie importante de ses ressources sur le centre (notamment le cluster de calcul, administré localement par un IE et un IR), pour lesquelles les utilisateurs et administrateurs ont besoin d'une grande adaptabilité et réactivité. Depuis, la politique des tutelles s'est assouplie, et s'avère compatible avec la démarche de BioSP. Néanmoins, nous regrettons que la construction d'une salle machine, initialement prévue dans le cadre de la construction du nouveau cœur de centre PACA (CPER 3A) n'ait finalement pas été retenue. L'unité devra veiller, en concertation avec ses tutelles et avec les autres unités du centre PACA, à disposer localement d'une salle informatique sécurisée face aux risques de pannes (actuellement, les systèmes d'ondulation et de climatisation donnent des signes de faiblesse ; bien qu'aucune interruption n'ait été à déplorer depuis 1 an, la climatisation a connu deux incidents majeurs en 2020 ayant nécessité une évolution matérielle en urgence au détriment d'une réflexion de long terme) tout en continuant à se diriger vers un fonctionnement hybride entre serveurs locaux / sur data-centers.

Domaine d'évaluation 2 : Attractivité

Référence 1. L'unité est attractive par son rayonnement scientifique et contribue à la construction de l'espace européen de la recherche.

Les membres de l'unité présentent régulièrement leurs travaux dans des congrès internationaux (plus de 145 présentations sur la période, notamment sur invitation à Spatial Statistics 2017, Lancaster [CNF-O17a]; Extreme-Value Analysis 2017 [CNF-O17], 2019 [CNF-O19], 2021 [CNF-O21] ; 5 conférences au CIRM ; Joint Meeting Brazil-France in Mathematics 2019 [CNF-R19a] ; 2 invitations au BIRS en 2016 [CNF-C16a,b] ; International Conference on Biomathematics and Biostatistics en 2019 [CNF-R19b]...). L'unité a organisé 9 congrès internationaux importants, dans différentes disciplines : en statistique notamment, avec la co-organisation des Journées de Statistiques (JdS) en compagnie du LMA d'Avignon Université, en EDP/Probas avec l'organisation d'événements au CIRM et en épidémiologie également avec notamment le séminaire final du projet Innov'Api qui a rassemblé plus de 1000 participants. Les chercheurs de BioSP ont

été impliqués significativement dans quatre réseaux internationaux (IRN ReadNet France-Japon-Corée-Taiwan sur le thème des équations de réaction-diffusion, LIA France-Chine sur les espèces invasives, réseau de collaboration INRAE-CSIRO sur la durabilité des résistances, la série de workshops SSIAB en statistique spatiale). BioSP est impliquée dans des activités éditoriales, au travers d'une participation au comité éditorial de 3 revues de statistiques. Deux membres de l'unité ont des responsabilités dans des sociétés savantes (deux membres élus dans le Groupe Environnement et Statistique de la SFdS). Deux chercheurs ont obtenu des prix nationaux, Thomas Opitz avec le Laurier Espoir Scientifique d'INRAE et Samuel Soubeyrand avec le Prix Jean Dufrenoy de l'Académie d'Agriculture de France. Trois prix de thèses ont été décernés aux étudiants de BioSP pendant la période.

Par ailleurs, BioSP a porté ou co-animé 4 « réseaux d'animation scientifique » sur la période. Ces réseaux sont des structures relativement pérennes, renouvelables sur une base annuelle ou bi-annuelle, ce qui a été le cas depuis au moins 6 ans pour les quatre réseaux en question. L'appel interne à INRAE (effectué via les départements de tutelle) « réseaux d'animation scientifique » a pour objectif de favoriser les échanges et susciter les collaborations entre équipes INRAE appartenant éventuellement à différents départements ou extérieures à INRAE, sur des thèmes méthodologiques et thématiques prioritaires pour le département finançant le réseau. La présence de porteurs de réseaux indique un dynamisme important des thématiques associées dans l'unité, et lui confère un certain rayonnement grâce à la régularité des événements organisés (en moyenne 3 événements par réseau et par an). Les réseaux co-animés par BioSP sont adossés à des priorités scientifique de l'unité (réseau RESSTE : priorité #Geolearning, réseau ModStatSAP : priorité #Epi-stat, réseau MEDIA : priorité #Evol-sys, réseau CiSStats : priorité #Obs-model). Les membres de l'unité participent également à des GdR, comme les GdR Ecostat et GeoSto dont les réunions 2019 à Avignon ont été organisées par l'unité ou encore le GdR MADICS.

Enfin, des membres de l'unité participent à des instances de pilotage de la recherche, le plus souvent internes à INRAE. Il s'agit notamment de participations au comité de pilotage de métaprogrammes (MP) INRAE (ACCAF ; CLIMAE ; MP en construction "XRISQUES"), aux conseils scientifiques des départements et même depuis septembre 2021 à la direction du département SPE (via un poste de chef de département adjoint), à des comités de suivi (convention cadre INRAE-Meteo-France ; initiative INRAE-Inria sur les risques naturels et environnementaux), à des ateliers de réflexion prospective (ARP Agroécologie, voir la synthèse [CHA-GLM20] ; Réflexion prospective pour les approches prédictives pour la biologie et l'écologie, à l'origine du MP DIGIT-BIO ; Réflexion prospective interdisciplinaire sur les risques naturels, alimentaires et environnementaux). En dehors d'INRAE, nous sommes régulièrement sollicités pour l'évaluation de projets ANR et ponctuellement pour participer à des comités ANR (e.g., CE32) et nous participons au Grand Enjeu Transversal "Risques naturels et environnementaux" d'AllEnvi.

Référence 2. L'unité est attractive par la qualité de sa politique d'accueil des personnels.

Sur le plan informatique, les doctorants/post-doctorants disposent tous d'un poste de travail portable et éventuellement d'un poste de travail fixe si l'encadrant le juge nécessaire. Ils ont accès au cluster de calcul de l'unité et bénéficient du soutien du PRI et des ingénieurs informaticiens de l'unité, pour l'utilisation de leur poste de travail (notamment sauvegarde), du cluster de calcul ou d'outils de calcul scientifique et de visualisation des données (notamment MSE et l'outil d'hébergement des applications Web de l'unité, cf Dom 3 Ref 1). Ils sont généralement répartis dans des bureaux de 2 ou 3, qu'ils peuvent conserver tout au long de leur contrat.

Sur la majorité de la période évaluée, et avant la pandémie (donc sur la période 2016-2019), les agents de BioSP étaient présents quotidiennement (sauf s'ils étaient en mission), permettant des interactions quasi-quotidiennes entre les doctorants/post-doctorants et leurs encadrants. Les échanges entre doctorants et encadrants se sont poursuivis de façon satisfaisante pendant la période "Covid" (2020-2021), au cours de laquelle 5 soutenances ont eu lieu et aucun abandon de thèse n'est à déplorer. Depuis 2021, les agents INRAE sont fortement incités à télétravailler. Les étudiants définissent leurs jours de télétravail avec leurs encadrants leur permettant a minima d'être présents simultanément 3 jours par semaine (cf. Dom 1 Ref 3). On notera 25 co-publications entre les doctorants de BioSP (ceux hébergés à BioSP, mentionnés dans les données de caractérisation) et leurs encadrants sur la période évaluée pour un total d'environ 30 doctorant*année, soit

environ 1 article par étudiant et par an. Sur l'ensemble des 206 publications de l'unité, 89 sont co-signées par des doctorants (certains n'étant pas hébergés à BioSP, cf Annexe 1, ou dans certains cas pas co-encadrés par un agent BioSP, elles ne sont pas comptabilisées dans les 25 ci-dessus). Six étudiants ont à la fois démarré et soutenu leur thèse pendant la période évaluée ; ils ont en moyenne publié 2.8 articles pendant leur thèse. Sur les 10 étudiants hébergés à BioSP ayant soutenu leur thèse durant cette période, 3 sont actuellement post-doctorants, 2 sont titulaires de la fonction publique (professeur en CPGE, IR à l'INSERM), 3 ont un emploi permanent dans le privé, 1 est en recherche d'emploi et 1 a une situation inconnue.

La mutualisation d'une grande partie des ressources récurrentes de l'unité, et la participation d'un nombre important d'agents à des projets dont l'unité est partenaire permettent à l'ensemble des agents de financer leurs inscriptions à des conférences internationales ainsi que leurs déplacements. Les étudiants sont encouragés à présenter leurs résultats dans le cadre du séminaire d'unité et dans des conférences.

L'unité a recruté 4 chercheurs juniors (CRCN) sur la période, dont 3 sur profil et 1 sur projet (concours très sélectif : 10 CRCN recrutés à l'échelle de l'institut pour 250 dossiers), ainsi qu'une ingénieure de recherche. L'unité a recruté une chercheuse senior (DR) sur concours externe. Parallèlement à ces recrutements au sein de l'équipe Recherche, l'unité a recruté 7 ingénieurs dans le cadre du montage de l'équipe OPE, et 2 techniciens (une GU et un PRI), tous deux via une mobilité interne. Comme pour les doctorants/post-doctorants, l'unité fournit un poste de travail portable, et toutes les ressources demandées par les personnes recrutées concernant l'informatique personnelle et le mobilier. Parallèlement à cette prise en charge sur les ressources récurrentes de l'unité, les recrutements sont souvent portés par un chercheur senior disposant de ressources via des projets : R Forien et F Patout ont par exemple été associés dès leur arrivée au projet ANR RESISTE, D Martinetti au projet UE XF-ACTORS et L Benoit au projet BEYOND et à la chaire Geolearning. L'unité incite (notamment via des discussions en réunion d'unité) et accompagne les jeunes chercheurs dans le dépôt de projets auprès des départements MathNum et SPE, voire auprès de l'ANR.

Sur le plan scientifique, les recrutements sont l'aboutissement d'un long processus. A l'exception du concours sur projet, les postes sont habituellement discutés collectivement puis portés par 1 ou 2 responsables ; l'arbitrage favorable par les tutelles (départements puis DG) n'intervient généralement qu'après quelques années, et une condition nécessaire pour cet arbitrage est précisément l'existence d'un contexte favorable à l'épanouissement scientifique du jeune recruté. Concernant le concours sur projet, le candidat a contacté l'unité pour construire ensemble le projet scientifique. Les jeunes recrutés sont donc naturellement associés aux travaux des porteurs dès leur arrivée dans le laboratoire et sont invités à participer aux différentes animations. La double structuration de l'unité, en axes disciplinaires et priorités scientifiques (cf. **Structuration et thématiques scientifiques de l'unité, Tableau 1 et Fig. 3** du Dom 3 Ref 1 ci-dessous) est particulièrement favorable en permettant aux agents d'identifier des interlocuteurs partageant leurs questions scientifiques sans nécessairement posséder la même formation scientifique. Les réseaux méthodologiques participent également à la bonne intégration des nouveaux arrivants à INRAE (au-delà de BioSP), en favorisant les collaborations hors unité (cf. Dom 2 Ref 1 pour une description des réseaux).

L'unité a invité trois chercheurs sur des périodes longues via des accueils en délégation : Edith Gabriel (statisticienne, 2014-2017), Florent Bonneau (statisticien, MCF à Avignon, 2019-2021) et Matthieu Alfaro (EDPiste, PR à Rouen, 2019-2021). L'accueil en délégation d'Edith Gabriel a permis de créer les conditions favorables à son recrutement comme DR dans l'unité BioSP en 2020 (notons que le recrutement de DR extérieurs à INRAE est assez exceptionnel, de l'ordre de 3 ou 4 postes sur 40 postes de DR proposés annuellement au concours). BioSP a accueilli de janvier à juin 2016 le séjour sabbatique de Denis Marcotte, professeur à Polytechnique Montréal. Ce séjour a abouti à deux publications communes [PUB-MA18a,PUB-MA18b].

Concernant les aspects autour de l'intégrité scientifique et de la science ouverte, cf le Dom 3 Ref 3.

Référence 3. L'unité est attractive par la reconnaissance que lui confèrent ses succès à des appels à projets compétitifs.

L'unité dispose d'importants financements via des appels à projets compétitifs, pour une moyenne de 306k€/an sur la période, à comparer aux 142k€/an correspondant aux financements récurrents. Elle répond avec succès

à des appels à projets internationaux (contrat européen INNOV'API "Innovation sanitaire pour la durabilité des exploitations apicoles" porté par l'unité) et des projets financés par les PIA (projet BEYOND, co-porté par l'unité et traitant en partie des méthodes de modélisation et d'analyse statistique pour améliorer l'épidémiologie et la prophylaxie en santé végétale) ou par France Relance (projet SEPIM, porté par l'unité et visant à améliorer, via l'analyse de données et la modélisation, la surveillance, l'évaluation, la prévision, l'interpolation et la mitigation des risques relatifs à la jaunisse de la betterave). Sur la période, 2 projets ANR portés ou co-portés par l'unité ont été acceptés : le projet SMITID (où il s'agit de proposer des méthodes statistiques pour inférer les chaînes de transmissions au cours d'épidémies sur la base de données de séquençage haut-débit) et le projet RESISTE (portant sur les aspects expérimentaux, théoriques et mathématiques liés à la question du sauvetage évolutif en partenariat avec l'Institut des Sciences de l'Evolution de Montpellier).

Parmi les financements dits "récurrents", certains sont eux-mêmes obtenus via des appels d'offres internes à INRAE, cf. Dom 1 Ref. 1. Les départements INRAE MathNum et SPE ont ainsi financé, via des appels d'offres internes, quatre $\frac{1}{2}$ thèses sur la période 2016-2021, l'autre $\frac{1}{2}$ financement étant obtenu via la Région PACA. Deux thèses ont été financées entièrement par l'ANR, deux autres actuellement en cours sont également financées par l'ANR dont une $\frac{1}{2}$ ANR et $\frac{1}{2}$ Métaprogramme.

Sur la période 2016-2021, l'unité a financé sur ressources propres un total de huit contrats doctoraux, cinq contrats post-doctoraux (soit 115 mois sur ressources propres contractuelles) et huit CDD "ingénieur" (soit 84 mois sur ressources propres contractuelles). Le cluster d'unité est majoritairement financé sur ressources propres. Cela correspond à un investissement de 74000€ sur la période évaluée, auxquels il faut ajouter environ 5000€/an de consommation électrique (sans prendre en compte le refroidissement de la salle machine).

Référence 4. L'unité est attractive par la qualité de ses équipements lourds et de ses compétences technologiques.

L'unité ne dispose pas d'équipement lourd bénéficiant d'un label. En revanche, il nous semble opportun de mentionner ici le cluster de calcul de l'unité, qui a contribué significativement à notre production durant la période (au moins 39 publications de l'unité se basent sur une utilisation significative du cluster de BioSP sur la période évaluée). Le cluster de calcul de BioSP a pour but de promouvoir le calcul scientifique à haute performance et permet l'accès à un environnement de logiciels scientifiques sous licence ou relevant du logiciel libre. Son parc de serveurs est constitué de 27 serveurs dédiés aux calculs soit 544 coeurs, 5112 Go de RAM. Quatre types de serveurs de stockage sont utilisés pour un total de 40 To. Cette plateforme de calcul destinée à la base aux personnels de l'unité BioSP s'ouvre à présent aux autres unités des Centres INRAE désirant effectuer des calculs numériques nécessitant des ressources informatiques importantes. Ces 4,5 millions d'heures théoriques de calcul annuelles ont été partagées ces dernières années par 133 utilisateurs différents dont plus d'un tiers issus d'autres unités essentiellement du centre PACA. Ces unités contribuent occasionnellement à l'achat de serveurs intégrés au cluster. L'unité dispose également d'une plateforme de virtualisation accolée au cluster de calcul et composée de 4 serveurs physiques, 128 coeurs (531 Go RAM, 121 To de disques), hébergeant 77 serveurs virtuels : 30 serveurs actifs (projets scientifiques, SI de BioSP) et 47 serveurs de test et de sauvegarde.

Synthèse de l'autoévaluation (Domaine 2)

Forces. L'unité BioSP est visible sur chacun de ses trois axes disciplinaires, malgré leur taille modeste. Nous avons organisé des événements internationaux sur chacun de ces axes, et les invitations régulières des scientifiques de l'unité à donner des conférences/séminaires/mini-cours indiquent une bonne reconnaissance académique. L'unité bénéficie par ailleurs d'une très bonne visibilité en interne (i.e., au sein d'INRAE) ; sa croissance importante durant la période évaluée (15 agents permanents recrutés dont 1 DR / 4 CR / 4 IR / 4 IE / 2 TR) et l'accueil d'enseignants-chercheurs en délégation sont de bons indicateurs de la confiance que lui accordent ses tutelles et de son attractivité. Ses succès aux appels à projets internationaux/nationaux/ internes lui confèrent des ressources propres lui permettant de fonctionner confortablement, de s'équiper de matériel informatique performant et de recruter régulièrement des doctorants et post-doctorants. Le dynamisme

scientifique de l'unité (illustrée notamment par les réseaux qu'elle anime), sa structuration croisée en disciplines et priorités scientifiques qui témoigne d'une perméabilité importante entre disciplines et les possibilités de collaborations sur le centre PACA offrent un cadre stimulant et attractif pour les jeunes recrutés (auquel contribuent les 2700 heures d'ensoleillement par an !).

Faiblesses. Un premier point de vigilance concerne l'accueil de doctorants. Sur la période, l'unité a accueilli 17 doctorants dans ses locaux. Certaines thèses ayant démarré avant la période, d'autres à la fin de la période, cela correspond à un total de 30 doctorant*année soit en moyenne 5 doctorants simultanément dans l'unité sur cette période de 6 ans. En comparaison d'une moyenne de 13 chercheurs sur la période évaluée, on arrive à un ratio doctorant/chercheur de 0.4 qui reste supérieur à la moyenne INRAE, qui est d'environ 1/4. Les 8 doctorants co-encadrés par des chercheurs de l'unité mais qui n'ont pas effectué leur thèse à BioSP ne sont pas comptés dans nos effectifs et n'apparaissent donc pas dans les données de caractérisation (ils sont donnés en Annexe 1). De façon générale, si des recrutements d'étudiants de bon niveau ont été possibles au cours de la période évaluée, les chercheurs de BioSP doivent se montrer proactifs pour les attirer : l'environnement universitaire de proximité (Université d'Avignon, dont le vivier d'étudiants en mathématiques est très limité), l'absence d'ED au sein d'INRAE et les modes de financement dont nous disposons peuvent rendre cette tâche complexe. L'unité n'étant pas adossée à un "gros" master, elle ne bénéficie pas d'un flux continu de bons étudiants et diversifie donc ses sources. Certaines possibilités de financement dont nous disposons (1/2 thèses départements) nécessitent une anticipation importante (presque 1 an) ne permettant pas de proposer une thèse "à la volée" lorsque nous rencontrons un bon étudiant de master et nécessitent un travail supplémentaire de recherche de l'autre 1/2 financement. Par ailleurs, certains financements de contrats doctoraux (notamment ceux des ENS, 2 financements durant la période) ne peuvent être versés qu'à des organismes possédant une école doctorale et l'accueil des étudiants doit dans ces cas passer par des conventions complexes avec l'université d'inscription (AMU dans ces deux cas).

Un second point de vigilance concerne le positionnement du curseur entre activités de recherche sur nos sujets propres et activités d'expertise et de conseil. La visibilité et la nature transversale des activités de BioSP sont en effet porteuses de sollicitations régulières pour participer à des projets appliqués (e.g. projets ANR portés par des biologistes) ainsi qu'à des instances de pilotage de la recherche. L'unité doit veiller à conserver un bon équilibre entre ses thèmes de recherche propres et ces différentes activités. En particulier, nous évitons de trop solliciter nos jeunes chercheurs et ingénieurs sur des missions de pilotage et d'organisation de la recherche.

Domaine d'évaluation 3 : Production scientifique

Référence 2. La production scientifique est proportionnée au potentiel de recherche de l'unité et répartie entre ses personnels.

Unité BioSP (les 2 équipes)

Les deux équipes Recherche et OPE composant l'unité BioSP ont des objectifs très différents en termes de production scientifique. Conformément aux profils d'activité présentés au début de ce document, la recherche est la principale activité de l'équipe Recherche, avec une valorisation principalement sous forme de publications scientifiques. L'activité de l'équipe OPE se concentre quant à elle sur de l'expertise technique. Les deux équipes interagissent néanmoins très régulièrement, notamment sur les enjeux correspondant à la Priorité Scientifique #Epi-stat et sur les métiers de l'informatique. Ainsi, les deux équipes participent ensemble au projet PIA4 BEYOND (qui vise une épidémiosurveillance de nouvelle génération), au projet INRAE TIERS-ESV (qui cible des avancées méthodologiques en fouille de texte pour la veille sanitaire), au projet INRAE SK8 (pour l'hébergement et la mise à disposition d'applications web) ou encore au projet "Plan National Dépérissement du Vignoble" CLIMESCA (concernant la structuration des données et l'analyse des facteurs de risque pour la gestion de l'ESCA, maladie de la vigne ; début en oct. 2022). Des articles de recherche sont en cours de préparation et pourront être comptabilisés lors de la prochaine évaluation.

Nous avons choisi de présenter les principales activités de l'équipe Recherche via le Dom 3 Ref 1 ci-dessous, et celles de l'équipe OPE via le Dom 4. Ces deux parties, qui correspondent respectivement aux activités cœur de métier de chaque équipe, sont volontairement hypertrophiées par rapport aux autres éléments de ce rapport.

Equipe n° 1 (Recherche)

Référence 1. La production scientifique de l'équipe satisfait à des critères de qualité.

Ainsi que nous l'avons discuté dans la présentation de l'unité, BioSP est structurée scientifiquement autour de 3 axes disciplinaires et 5 priorités scientifiques. Nous détaillons en **Fig. 3** les sous-domaines principaux associés aux trois axes disciplinaires ainsi que les interactions entre axes disciplinaires et priorités scientifiques.

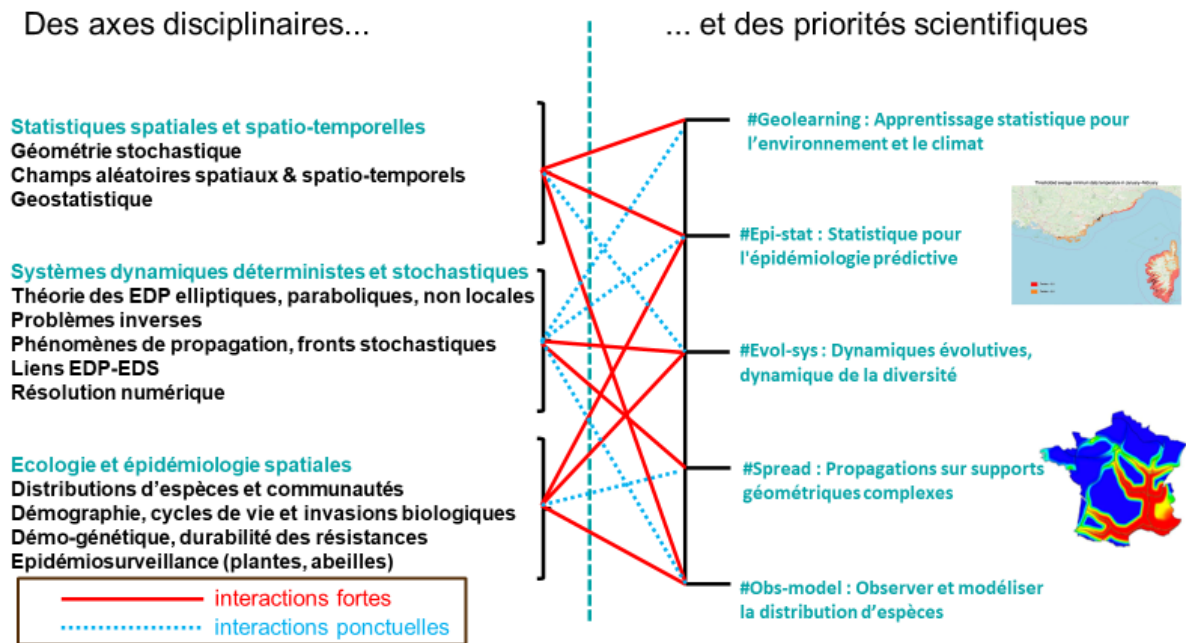


Fig. 3. Détail des axes disciplinaires et interactions avec les priorités scientifiques.

Ci-dessous, nous avons choisi de présenter la production scientifique de l'équipe au travers des cinq priorités scientifiques, puis d'une section spécifique sur nos travaux en rapport avec l'épidémie de Covid-19 et d'une section présentant des éléments additionnels de projet.

Apprentissage statistique pour l'environnement et le climat

En géostatistique et statistique spatiale, les recherches récentes portent sur des avancées théoriques et méthodologiques et des applications diverses. De nombreux résultats théoriques ont été obtenus, le plus souvent par le biais de collaborations internationales, et en privilégiant des journaux à forte notoriété. Ainsi, en théorie des champs aléatoires, nous avons: établi une caractérisation complète de l'anisotropie des champs aléatoires spatiaux [PUB-ASP16], proposé une nouvelle classe flexible de fonctions de covariance croisée pour la modélisation de données spatio-temporelles multivariées [PUB-BAP16], fait des avancées théoriques pour l'analyse des données de composition spatiale [PUB-AM16, PUB-CAM21]. Des liens fructueux entre statistiques spatiales et physique statistique conduisant à une nouvelle famille de champs aléatoires de Boltzmann-Gibbs [PUB-AHO21] ont été établis, et des avancées significatives pour la simulation de champs aléatoires spatio-temporels [PUB-AEL20, PUB-CAP20] et des approches de réduction pour les covariances non stationnaires [PUB-PSM20, PUB-SP21] ont été proposés. Concernant les processus ponctuels, des méthodes de prédiction de l'intensité en dehors de la fenêtre d'observation ont été proposées [PUB-GCC17, PUB-GBM16], et de nouveaux modèles spatio-temporels multi-échelles ont été obtenus [PUB-RBG20, PUB-RBG21]. L'approche EDPS (pour Equations aux Dérivées Partielles Stochastiques) proposée dans Lindgren et

al. (2011) est un changement de paradigme en statistiques spatiales, permettant d'aborder des ensembles de données de très grande taille. Nous avons obtenu de nouveaux résultats théoriques pour l'extension de cette approche à la modélisation de données spatio-temporelles [PUB-CAD21], incluant aussi les termes de transport et de réaction [PUB-RAS21]. Une thèse de doctorat (Lucia Clarotto, cf Annexe 1) a débuté en 2020 afin de proposer des méthodes et des algorithmes capables de considérer de très grands ensembles de données spatio-temporelles.

Les applications sont nombreuses, touchant à la mesure LIDAR en foresterie [PUB-PAS18, PUB-SPA20], l'agriculture numérique [PUB-LBA17a, PUB-LBA17b], la santé des forêts [PUB-CKA17, PUB-TBA20] et les variables climatiques [PUB-BAP16, PUB-BAM18], l'analyse des adventices [PUB-GCC17, PUB-GBM16], les feux de forêt [PUB-GBO17, PUB-OBG20] et les mouvements de terrain [PUB-LOA20].

A la croisée de la théorie des valeurs extrêmes et des statistiques spatiales, l'analyse des événements extrêmes ayant une extension spatiale et temporelle est un domaine de recherche très actif auquel BioSP contribue significativement, notamment en lien avec le changement climatique. Nos contributions couvrent à la fois des aspects théoriques et leurs applications dans les domaines des risques climatiques et environnementaux, y compris les aspects suivants :

- des résultats théoriques sur la caractérisation des extrêmes conjoints (multivariés, spatiaux, temporels), souvent basés sur des variables latentes (i.e. cachées) et utilisés pour modéliser des chocs extrêmes agissant sur plusieurs sites ou composantes [PUB-HOT17, PUB-BGO19, PUB-EOW19, PUB-HOT21, PUB-ZHO22a] ;
- des modélisations originales pour tenir compte des tendances et des corrélations non stationnaires des extrêmes, par exemple en contexte de changement climatique [PUB-OHB18, PUB-ZHO22b] ;
- la simulation stochastique hybride d'événements extrêmes combinant la théorie des valeurs extrêmes pour les processus stochastiques et l'apprentissage automatique [PUB-OAM21, PUB-PTC20] ;
- la modélisation spatio-temporelle de motifs ponctuels représentant des événements extrêmes environnementaux [PUB-PFO21].

Ces méthodes ont été appliquées aux processus météorologiques et climatiques, souvent dans le but de saisir les effets du changement climatique, par ex. pour les précipitations [PUB-OHB18, PUB-BGO19, PUB-PTC20], les vitesses de vent [PUB-MOL17, PUB-HOT17, PUB-ZHO22a] ou des vagues de chaleur [PUB-OAM20, PUB-ZHO22b], et aux risques environnementaux tels que les incendies de forêt [PUB-PFO21].

L'investissement de BioSP dans le champ de l'apprentissage automatique et du machine learning est assez récent, ce qui entraîne que les contributions se limitent pour l'instant à des applications dans le champ de l'épidémiologie végétale [PUB-MS19, PUB-TB20]. De nombreux projets ont récemment démarré, et ceux-ci n'ont pas encore abouti à des publications. Depuis plusieurs années, l'unité considère que la prise en compte des dépendances spatiales dans les méthodes d'apprentissage est une priorité scientifique. L'unité porte depuis 2019 un projet de recrutement CRCN "Méthodes d'apprentissage en statistiques spatiales".

Note : Thomas Opitz a reçu le "Laurier Espoir Scientifique" 2020 attribué par INRAE, en partie pour ses contributions de premier ordre dans ce domaine, cf Dom 2 Ref 1.

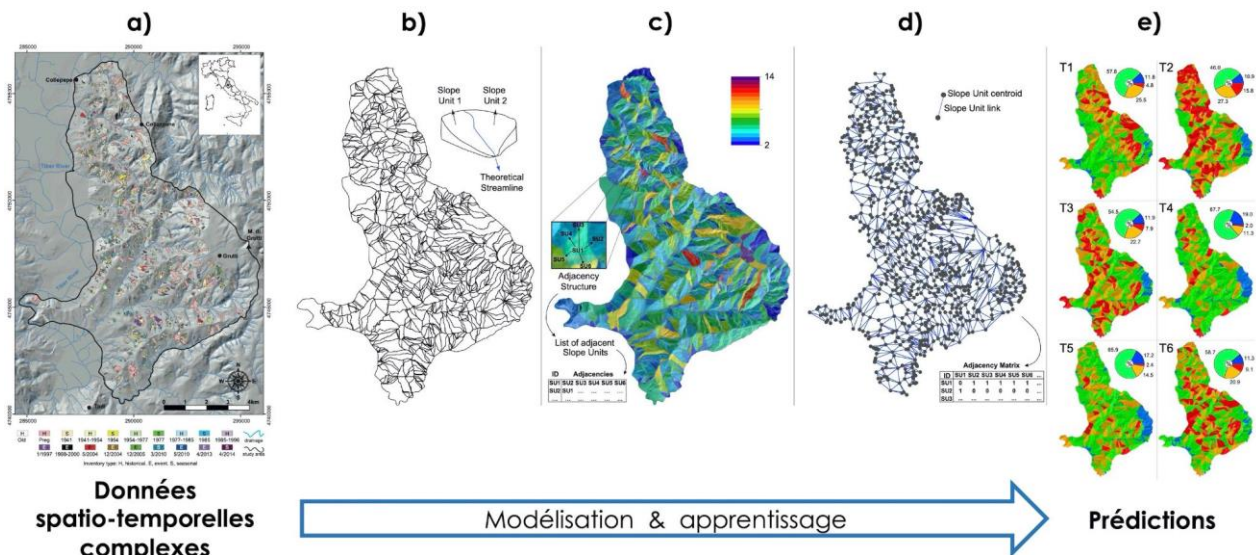


Fig. 4. Illustration d'une approche de modélisation statistique spatio-temporelle pour des données d'occurrences de mouvements de terrain : domaine d'étude et observations discrètes (a) ; maillage spatial adapté au processus observé, ici des "slope units" = versants (b) ; voisinages spatiaux pour définir un modèle d'autocorrélation (c) et sa représentation sous forme de réseau (d) ; cartes de prédictions probabilistes spatio-temporelles (e).

Statistique pour l'épidémiologie prédictive

A travers cette priorité, nous menons des recherches sur les approches prédictives en épidémiologie (e.g. prédire la dynamique future d'une épidémie, prédire l'efficacité d'une méthode de surveillance, prédire les conditions des contournements de résistances, prédire l'impact de polluants sur la santé de populations...) en mobilisant des données hétérogènes, la modélisation, la statistique, l'analyse de modèles et les développements informatiques. Une partie des sujets de recherche incorporés à cette priorité visent à comprendre les dynamiques épidémiques plus qu'à prédire, car le travail de compréhension est un travail « amont » généralement nécessaire à la prédiction. Après sa mise en place en 2017, la priorité a bénéficié du recrutement d'un CR, Davide Martinetti en 2018, a porté en partie l'interaction avec l'équipe OPE qui compte notamment dans ses rangs deux ingénieurs en statistique (Lucie Michel et Marine Marjou), s'est concrétisée dans plusieurs projets (UE : XF-ACTORS, IPM-Popillia ; EFSA : BriskaR-NTL ; ANR : SMITID*, SPREE ; Plans nationaux : BEYOND*, SEPIM*, Co-ACT2 ; INRAE : PHYTOSENTINEL*, INITIATE, TIERS-ESV* ; * : projets (co-)portés par l'unité qui ont en particulier permis de recruter 2 post-docs (V Baudrot, M Ribaud), 5 doctorants (C Abboud, M Alamil, M Choufany, B Vieille, A Radici) et 2 ingénieurs d'étude (D Chauvin, M Siegwart), et a été récompensée par la remise de trois prix individuels (2 prix de thèse et 1 prix de l'académie d'agriculture pour un chercheur permanent). L'animation de la priorité, au-delà de discussions et collaborations internes, a été en partie adossée au réseau ModStatSAP (~160 inscrits à la liste de diffusion ; 10 symposiums sur la période évaluée) auquel nombre d'agents de BioSP participent régulièrement.

Une partie des recherches menées, impulsée notamment par l'arrivée de Davide Martinetti, concerne la construction et l'analyse de **réseaux spatiaux et spatio-temporels pour l'épidémiologie**. Ces réseaux servent notamment à appréhender la propagation des pathogènes des plantes portés par les mouvements des masses d'air [PUB-LMC18,PUB-CMS21a,PUB-CMS21b,Choufany2020(thèse)] et à prédire l'efficacité de stratégies améliorées de surveillance des épidémies [PUB-MS19,Choufany(2020,thèse)]. Un autre axe de la priorité concerne les **approches mécanistico-statistiques** (associant un modèle mécaniste du phénomène étudié, un modèle probabiliste du processus d'observation et une méthode d'inférence des paramètres et des processus cachés, cf. Soubeyrand & Roques 2014) **adaptées aux dynamiques de pathogènes** [PUB-SDM18,PUB-PSD18,PUB-ABP19, PUB-PBW21,PUB-NTS21,Abboud(2019,thèse)]. Ce cadre repose notamment sur

l'adaptation de méthodes d'inférence telles que des algorithmes Monte Carlo adaptatifs mais aussi sur la formalisation de modèles originaux. A cet égard, des recherches ont été menées spécifiquement sur la **représentation et l'analyse mathématique de la dispersion de groupes dans les modèles de propagation**, en utilisant des processus de points ou d'objets [PUB-MS17,PUB-DM19] ainsi que des marches aléatoires markoviennes [PUB-SL17]. De plus, de nouveaux modèles couplant micro-évolution des pathogènes et dynamique épidémique ont été proposés pour la **reconstruction probabiliste des transmissions inter-individuelles sur la base de données génomiques** [PUB-S16,PUB-AHB19,Alamil(2020,thèse)]. S'agissant toujours de modélisation, des modèles multi-échelles fondés sur des représentations explicites de l'espace ont été développés pour produire des outils d'aide à l'**évaluation spatio-temporelle des risques écotoxicologiques**, avec comme cas d'étude le risque létal engendré par le pollen de maïs génétiquement modifié sur des espèces de lépidoptères non cibles [PUB-LWMS18,PUB-WLR19,PUB-BWL21]. On retrouve également dans la priorité le développement de **méthodes statistiques pour l'estimation des traits d'histoire de vie de pathogènes sur la base de données génétiques** [PUB-SSG16,PUB-RMM17,PUB-SGM17,PUB-EBS21]. Une autre dimension de la priorité concerne l'**adaptation des méthodes d'analyse de sensibilité aux modèles spatio-temporels d'épidémies** afin de comprendre les paramètres clefs des épidémies [PUB-MFC20] ou des stratégies de surveillance [PUB-RBD18,PUB-RBD19,PUB-PSJ19]. Enfin, les travaux menés sur le **Covid-19**, qui font l'objet plus bas d'un paragraphe spécifique, contribuent également à la priorité. Une partie des travaux mentionnés ici reposent sur des **apports importants de l'ingénierie de l'unité** ou sont valorisées par celle-ci. Ces apports concernent notamment la production de packages R (e.g., briskaR [LOG-BRL21], StrainRanking [LOG-MG21], FeedbackTS [LOG-SMB20]), la mise à disposition d'applications Web permettant la visualisation dynamique des données et des sorties de modèles et analyses statistiques, le partage de données et l'utilisation de l'environnement de calcul MSE dont il est question plus bas.

La priorité « Statistique pour l'épidémiologie prédictive » s'appuie sur une forte dynamique scientifique et des enjeux sociétaux majeurs, comme illustré par la création de la Plateforme ESV ou les crises *Xylella fastidiosa* et Covid-19. Au-delà des travaux marquants en cours, comme ceux concernant nos apports méthodologiques dans le champ de l'épidémiologie, nous envisageons d'étendre nos recherches selon deux axes décrits plus longuement dans le dernier paragraphe de cette section : les approches statistiques intégrant des données hétérogènes de diversité et permettant de comprendre et prédire les dynamiques spatio-temporelles et multi-échelles de communautés telles que les phytobiomes ; le développement d'approches hybridant apprentissage profond et modèle mécanistico-statistique pour incorporer dans nos études des données massives informant les dynamiques écologiques et épidémiologiques à de grandes échelles spatiales.

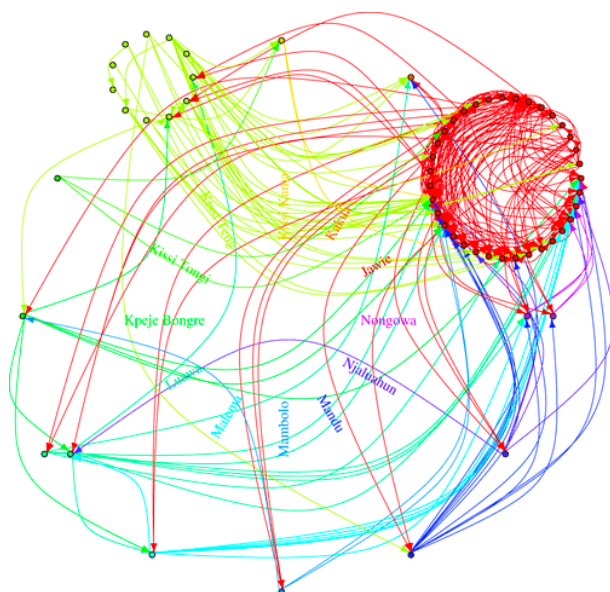


Fig. 5. Reconstruction (à partir de données génomiques caractérisant la diversité intra-hôte du pathogène et à l'aide d'une méthode d'apprentissage) des liens épidémiologiques les plus probables entre patients atteints de la maladie Ebola en Sierra Leone [PUB-AHB19].

Dynamiques évolutives, dynamique de la diversité

Etude d'EDP et d'équations intégro-différentielles décrivant des populations structurées en traits. Ce thème a émergé dans l'unité à partir de 2015 suite à une première collaboration [PUB-MR16] avec un biologiste de l'Institut des Sciences de l'évolution de Montpellier (UMR 5554), laboratoire reconnu en sciences de l'évolution. Depuis, nos recherches sur les systèmes dynamiques ont connu une forte inflexion vers des questionnements issus de la biologie évolutive, ce qui a conduit à la création de la priorité scientifique « Evol-Sys ». Ce thème permet de mobiliser des recherches sur des fronts de sciences en mathématiques (EDP non locales non linéaires [PUB-HLM20,PUB-AGH21], équations intégro-différentielles avec noyau hétérogène [PUB-GHM19], étude de valeurs propres pour des systèmes elliptiques avec coefficients non constants et non bornés [PUB-HLR21], problèmes à frontière libre [PUB-RGM17], à l'intersection entre EDP et systèmes dynamiques stochastiques [PUB-ALR19,PUB-LR21,PUB-PFA21], en collaboration avec des mathématiciens de premier plan comme F Hamel (I2M, Aix-Marseille Univ.), A Lambert (LPSM, Sorbonne Univ.) ou M Alfaro (LMRS, Univ. Rouen). Il est porteur d'enjeux sociétaux majeurs tels que l'émergence d'une antibiorésistance pour une bactérie ou l'apparition d'un changement d'hôte chez un virus (sujet d'actualité ...). Ces enjeux sont centraux pour INRAE, et ont permis d'entamer de nouvelles collaborations avec des biologistes d'INRAE PACA et Nouvelle-Aquitaine (unités Pathologie Végétale et SAVE) via plusieurs projets ANR et Ecophyto (ARCHIV, RESISTE et MEDEE, cf ci-dessous). Au-delà d'INRAE, cette thématique intervient fortement dans la collaboration avec le CSIRO de Canberra, qui a abouti à plusieurs publications étudiant des stratégies de déploiement de la diversité génétique au sein de paysages agricoles afin de limiter les capacités évolutives des agents pathogènes [PUB-BZB16, PUB-RPB18,PUB-PRB18,PUB-RPR18,PUB-RFP21] ainsi qu'à la production d'un logiciel [LOG-RRP21]. Nos recherches sont valorisées par des articles dans des revues de qualité, tant en mathématiques (*Analysis and PDE, Journal de Mathématiques Pures et Appliquées, SIAM Journal on Applied Mathematics, Nonlinearity, ...*) qu'en biologie évolutive (*Genetics, Evolution, Evolutionary Applications ...*) sans négliger les revues/processus de science ouverte (*PCI Evol Biol, PCI Math & Comp Biol*). Une partie de ces travaux s'est appuyée sur un espace de travail collaboratif Matlab/JupyterHub mis en place par un ingénieur de l'unité.

Pour répondre aux objectifs des biologistes, nous avons introduit des méthodes nouvelles d'analyse d'EDP et d'équations intégro-différentielles. Les résultats classiques en temps long ne sont en effet pas adaptés à la compréhension des phénomènes de sauvetage évolutif ou d'extinction. Il s'agissait donc de proposer des méthodes conduisant à des résultats quantitatifs et qualitatifs pendant les phases transitoires et stationnaires. Ces méthodes consistent principalement à définir des variables d'intérêt (typiquement la *fitness*, ou valeur sélective) et à déduire de l'équation initiale une équation sur la dynamique d'une fonction génératrice des moments associée à ces variables (cette méthode est bien décrite dans [PUB-HLM20]). Nous donnons un exemple dans le portfolio [PUB-RPB20], ici appliqué à des EDP décrivant des changements environnementaux (correspondant p ex à des scénarios thérapeutiques ou des phénomènes écologiques).

Au-delà des publications, cette dynamique forte dans l'unité s'est notamment matérialisée par le recrutement de deux CRCN sur des profils en lien avec ce thème (R Forien et F Patout), l'accueil en délégation d'un MCF sur 2 ans (M Alfaro), le co-encadrement réussi de 3 thèses (dont 2 en mathématiques : M-E Gil et F Lavigne et 1 en écologie : P Zamberletti) et de 2 post-docs (1 en mathématiques: F Patout et 1 en épidémiologie: M Zaffaroni) et l'élaboration du projet ANR RESISTE (PI : G Martin ISEM, co-PI : L Roques). En s'étant fortement investie sur ce thème actuellement moteur pour les maths-bio sur le plan national et international, l'unité BioSP a acquis une bonne visibilité, notamment via la co-organisation d'une semaine au CIRM sur ce thème en 2020 et de 6 journées scientifiques dans le cadre du réseau MEDIA (<https://media.biosp.org>) et de la FRUMAM. Les projets sont nombreux et mobilisent des compétences théoriques en EDP / probabilités / biologie évolutive ainsi que des compétences d'ingénierie en analyse numérique des EDP : couplage avec des EDS pour la quantification de probabilités de sauvetage évolutif, prise en compte d'hôtes multiples conduisant

à des problèmes d'optimisation de valeurs propres pour des systèmes, prise en compte de la variance entre trajectoires d'adaptation, analyse de phénomènes de propagation avec prise en compte de l'adaptation sur un ou plusieurs traits (comme dans [PUB-AGH21]), étude des lignées ancestrales, comparaison des résultats théoriques avec les résultats expérimentaux du projet RESISTE ...

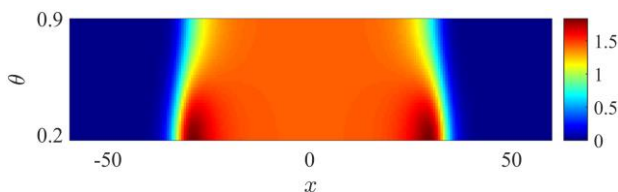


Fig. 6. Simulation d'une propagation spatiale (axe x) avec prise en compte de l'adaptation suivant un trait affectant la force de l'effet Allee (axe θ), d'après [PUB-AGH21].

Structure spatiale de la diversité génétique. Historiquement, les systèmes dynamiques étaient mobilisés à BioSP essentiellement dans le cadre de la compréhension de phénomènes de propagation *spatiale* avec notamment l'introduction de méthodes autour de la notion de dynamique interne d'un front de propagation (typiquement, les solutions de type "travelling wave" d'EDP de réaction-diffusion, que l'on qualifie de poussées ou tirées) (Garnier et al 2012, Roques et al 2012). Ces notions, qui trouvent des applications dans la compréhension de la dynamique diversité génétique neutre au cours d'une expansion, sont maintenant largement utilisées par les mathématiciens et les biologistes (voir par exemple les travaux en maths-bio de [Mark Lewis](#)). Nos recherches dans ce domaine sont toujours actives, avec l'étude de phénomènes de propagation en mathématiques [PUB-AC17,PUB-GHR17,PUB-AGH21,PUB-CGZ21] comme en écologie théorique [PUB-RWF16,PUB-HML21,PUB-MHL21,PUB-RDB21], ou même à partir de données d'expansions en cours [PUB-BXK6,PUB-GKJ17,PUB-LKR21,] et entrent en résonance avec les travaux récents de probabilistes proches de l'unité (Foutel-Rodier & Etheridge 2020). Des travaux sont également en cours avec A. Etheridge et S. Penington (Univ of Oxford) sur l'étude de solutions poussées dans un cadre stochastique. Nous avons encadré un stagiaire de M2 (Nathanaël Boutillon, ENS Lyon) sur ces thèmes ; il poursuivra une thèse à la frontière entre probabilités et EDP dès 2022 dans le cadre d'une collaboration BioSP/I2M.

Parallèlement à ces travaux, nous avons étudié les propriétés de processus stochastiques, et notamment leur convergence vers des extensions du mouvement brownien (biaisé, réfléchi) pour décrire les conséquences d'hétérogénéités spatiales [PUB-F19a] et l'impact d'une barrière à la dispersion [PUB-F19b] sur l'évolution de la composition génétique spatiale d'une population.

Propagations sur supports géométriques complexes

Cette priorité concerne pour l'essentiel la modélisation, l'analyse et la compréhension des phénomènes de propagation dans des paysages réalistes. Dans ce cadre, nos recherches ont porté principalement sur trois aspects :

- La poursuite de l'analyse et la simulation de systèmes d'EDP sur des supports couplant des éléments géométriques de dimensions différentes, initiées dans le travail [PUB-RB16] où l'unité a développé un formalisme mathématique et numérique pour la modélisation et la simulation d'EDP sur des géométries 1D-2D, couplant des dynamiques sur des éléments linéaires (1D) et des éléments bidimensionnels (2D). Cette approche permet notamment de modéliser des dynamiques de populations via des modèles EDP définis sur des zones d'études réalistes, de type paysages agricoles. Ces recherches ont pu être valorisées dans la thèse de P. Zamberletti [PUB-ZSO21] sur l'étude d'un modèle de dynamique de population novateur dans un paysage agricole aléatoire composé de parcelles et de haies.

- Le besoin d'ajuster les modèles EDP aux observations faites sur le terrain nous a conduit à développer un environnement informatique, MSE (pour Mechanistic-Statistical Environment) [LOG-BR16,LOG-BW18], dédié aux approches mécanistico-statistiques. Cet environnement est de plus en plus utilisé dans le cadre de collaborations variées [PUB-RB16,PUB-ABP19,EXP-BR19] et l'intégration des modèles 1D-2D est en cours. Le cœur de simulation des systèmes EDP est actuellement Freefem dont les limites sont atteintes. Afin de répondre à une demande croissante en interne et à l'extérieur de l'unité et de permettre la montée en puissance de MSE dans MathNum et à INRAE, il est devenu nécessaire de s'orienter vers des bibliothèques de simulation (Feel++, Dolfin). Il s'agira alors d'implémenter des méthodes numériques dans un langage compilé (C,C++) et de définir une interface programmatique rendant ces codes utilisables par les scientifiques et exploitables sur le cluster de l'unité. Pour réaliser ces missions, nous faisons depuis 2 ans une demande de recrutement d'un ingénieur (IE) ayant de solides compétences en algorithmique et en calcul scientifique (demande arbitrée favorablement par MathNum, mais non retenue par la DG en 2021). Dans l'attente d'un poste pérenne, l'unité a choisi de recruter un ingénieur en CDD (Tanguy Theulier).

- Le développement d'outils d'analyse d'équations intégrées-différentielles (IDE) en présence d'obstacles où l'on s'intéresse à l'étude analytique des solutions de types front généralisés dans un domaine perforé. Ces travaux sont au cœur de la thèse de J. Brasseur et ont permis des développements théoriques et numériques importants [PUB-BCH19,PUB-BC20,PUB-BC21] notamment dans la prise en compte de distance pertinente associée à la géométrie du domaine (voir simulation ci-contre avec distance géodésique vs euclidienne). Les résultats novateurs obtenus ont permis à J. Brasseur d'obtenir le prix de meilleure thèse de l'université d'Aix Marseille 2018. Ces travaux ouvrent de nombreuses perspectives notamment vis à vis de l'intégration de modèles IDE dans les modèles 2D-1D déjà développés dans l'unité.



Fig. 7. Dans le travail [PUB-BC21], des développements théoriques et numériques importants ont été réalisés sur les modèles de réaction diffusion non-locaux permettant la prise en compte de distances pertinentes associées à la géométrie du domaine (e.g. distance géodésique, euclidienne, ...).

Observer et modéliser la distribution d'espèces

Décrire et comprendre comment la biodiversité se structure dans l'espace et dans le temps est un préalable indispensable pour élucider les mécanismes majeurs gouvernant la distribution des espèces et mettre en place des mesures de gestion de la biodiversité. Dans cette priorité scientifique, nous menons des recherches dans trois directions. (1) si la plupart des modèles de distribution d'espèces (SDM) sont basés sur une approche corrélative, l'approche mécanistico-statistique nous permet d'aller au-delà en intégrant des mécanismes au sein des SDM [PUB-LPD20,PUB-PBW21,PUB-RDB21]. L'estimation des paramètres de tels modèles peut être coûteuse en temps de calcul lorsque l'on utilise des algorithmes classiques. Aussi, nous nous attachons au développement d'algorithmes efficaces, une perspective étant l'utilisation d'approches d'apprentissage profond (Physics Informed Deep Learning, cf "Éléments additionnels de projet", ci-dessous). Jusqu'à présent, aucun logiciel dédié à l'approche mécanistico-statistique n'existait ce qui rendait son transfert plus compliqué vers les écologues. Désormais le développement de l'environnement MSE par les ingénieurs de BioSP (PS Propagations sur supports géométriques complexes) comble cette lacune et son apport semble indéniable. (2) Si les SDM prennent généralement la dimension spatiale en compte, peu d'entre eux font une modélisation spatio-temporelle. L'approche mécanistico-statistique permet naturellement de le faire puisque les dynamiques sont modélisées de façon explicites. Une autre extension possible des SDM est, dans un cadre de régression, de considérer des champs spatio-temporels. (3) De plus, dès lors que l'on s'intéresse aux dynamiques de

communautés, la variable étudiée est multivariée. Ces développements pour l'analyse de données spatio-temporelles multivariées sont traités dans la thèse de F. Lasgorceux et étaient au cœur du projet INRAE INITIATE.

L'inférence des modèles développés dans cette PS repose généralement sur des données ayant une grande emprise spatiale et temporelle, et pouvant considérer également un grand nombre d'espèces. Sur de telles emprises, les données sont souvent complexes du fait de la multitude d'observateurs, de la diversité des protocoles d'observation (y compris l'absence de protocole), de leur quantité à la fois importante et faible au regard des processus à inférer et de leur nature spécifique-spatio-temporelle. Ainsi, leur analyse nécessite une attention toute particulière afin de séparer ce qui relève du processus d'acquisition des données de la dynamique écologique sous-jacente étudiée. Cette problématique a été étudiée lors de la thèse de C. Botella [PUB-BJB18,PUB-BJM20,PUB-BJB21] et est au centre de la thèse de F. Lasgorceux. Une particularité des travaux menés dans cette thématique est la volonté d'être au plus près des porteurs d'enjeux en se confrontant aux questionnements des gestionnaires d'espaces naturels. Ceci se traduit par plusieurs publications [PUB-BPA19,PUB-FWP21] et par une thèse en co-encadrement avec un parc national. Dans ce contexte, le partage des codes informatiques et de la chaîne de traitement des données est un point important menant à une collaboration étroite avec les ingénieurs informaticiens de BioSP. Enfin, à cette PS s'adosse le réseau INRAE CiSStats (Citizen Science Statistics) porté par BioSP qui vise à transférer des développements statistiques vers les porteurs d'enjeux.

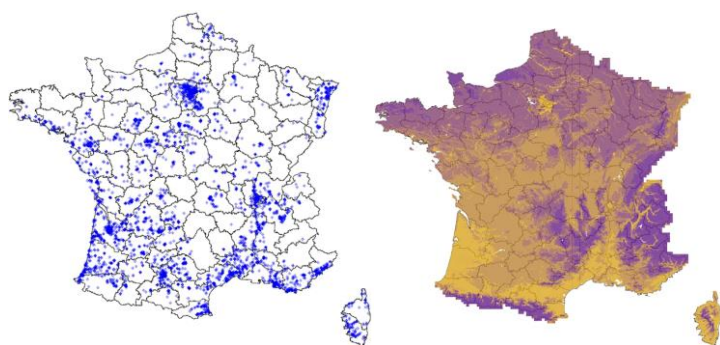


Fig. 8. La carte de gauche montre les données d'occurrence de *Phytolacca americana* observées par les utilisateurs de l'application *Pl@ntNet* avec une identification automatique de l'espèce pour les années 2017 et 2018 (4640 observations). La carte de droite montre la densité relative de cette plante estimée à partir de ces données grâce à la méthode développée dans [PUB-BJB21].

Modélisation de l'épidémie de Covid-19

Les recherches motivées par l'épidémie de Covid, si elles ne mobilisent pas toujours les outils les plus pointus que nous développons dans chaque Axe Disciplinaire, illustrent parfaitement l'une des spécificités de BioSP : la circulation des idées et les interactions fortes entre disciplines académiques différentes. En collaboration avec des mathématiciens de premier plan (Etienne Pardoux, I2M, Henri Berestycki, EHESS) et des médecins, l'équipe a publié 6 articles [PUB-RBB20, PUB-RKP20a, PUB-RKP20b, PUB-SDR20, PUB-SRB20, PUB-FPP21]. En s'appuyant sur l'approche mécanistico-statistique, via un couplage entre modèles d'EDO de type SIR, modèle probabiliste et approche statistique bayésienne, nous avons pu calculer le taux de létalité réel lié à la Covid-19 dès mars 2020 [PUB-RKP20a]. Il s'agissait alors d'une première mondiale en dehors de la Chine, relayée par INRAE via un [communiqué de presse](#). L'originalité sur le plan méthodologique venait ici du modèle d'observation, qui faisait intervenir la probabilité relative de subir un test pour un individu sain vs infectieux, et l'estimation de ce paramètre. Fin avril, les calculs de l'Institut Pasteur avaient corroboré ces résultats avec un modèle plus complexe et faisant appel à un jeu de données plus conséquent, incorporant des données issues de l'étranger.

Ces recherches ont également donné lieu à des développements méthodologiques originaux venant renforcer nos collaborations avec l'I2M dans des disciplines différentes (EDP via M Hauray, probabilités via E Pardoux, sciences des données via Denys Pommeret), avec un modèle de propagation spatio-temporelle non local [PUB-FPP21] qui a motivé un sujet de thèse à l'I2M, le développement et l'analyse de modèles probabilistes non-markoviens en collaboration avec l'I2M ainsi qu'une méthode orientée données pour la prédiction du nombre de morts sur un horizon de quelques semaines [PUB-SRB20]. Ces travaux illustrent également la bonne interaction entre recherche et ingénierie, avec la mise en place à la volée d'[applis web](#) [LOG-RSB20a, LOG-RSB20b, LOG-SR20] présentant de façon interactive certains de ces résultats, notamment une prédiction pays par pays du nombre de morts directement dus à la Covid-19 (d'après [PUB-SRB20]), le taux d'immunité local en France (d'après [PUB-RBB20]) et le suivi en temps réel du nombre de reproduction efficace R_t pays par pays (d'après [PUB-SDR20]). La nécessité de publier rapidement certains résultats (évaluation nécessaire par des pairs avant diffusion des résultats par INRAE, notamment sur le calcul du taux de létalité réel et le risque de seconde vague) et la difficulté que nous avons rencontrée à publier nos travaux jugés trop techniques par les revues médicales de premier plan, nous ont orientées dans certains cas vers des revues en "open access" à publication rapide, avec toutefois un processus d'évaluation tout à fait standard.

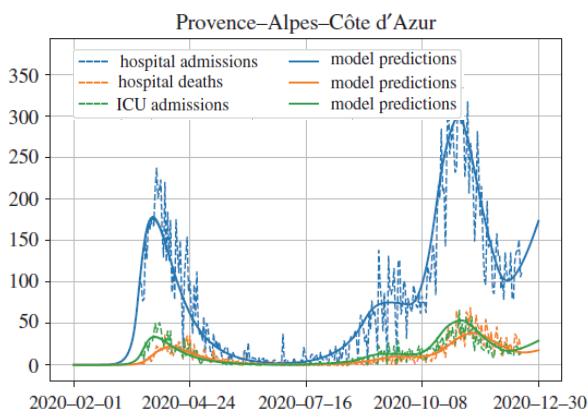


Fig. 9. Données hospitalières vs prédictions du modèle SEIR non markovien de [PUB-FPP21] en région PACA, obtenues après estimation de la distribution du temps écoulé entre l'infection d'un individu et son admission à l'hôpital ainsi que son décès (pour ceux qui décèdent à l'hôpital). La limite déterministe du modèle conduit à un système intégro-différentiel.

Elements additionnels de projet

L'essentiel du projet de l'équipe Recherche se situe dans la continuité de nos travaux et a été en partie décrit dans les items ci-dessus. Mentionnons par exemple les développements autour de l'approche EDPS en statistique spatiale ; la prise en compte des dépendances spatiales dans les méthodes d'apprentissage ; les développements méthodologiques dans le champ de l'épidémiologie ; le développement et l'analyse de modèles de biologie évolutive couplant méthodes d'EDP et de probabilités ; le développement d'algorithmes robustes en C et C++ pour MSE ; l'analyse de données de distributions d'espèces non nécessairement protocolées. Avec les recrutements récents de CR/DR (Davide Martinetti, Edith Gabriel, Raphaël Forien, Florian Patout, Lionel Benoît), l'équipe Recherche a renforcé son socle disciplinaire (en statistique, probabilité et analyse mathématique) au cours des 5 dernières années et est donc en mesure d'aborder sereinement la plupart de ces aspects.

Parallèlement à ces aspects, nous souhaitons investir dans le domaine de l'apprentissage automatique et de son hybridation avec nos approches historiques. Cette évolution est incarnée à travers différents éléments de projets venant contribuer aux différents axes disciplinaires et aux priorités scientifiques actuelles.

- Il s'agit en particulier de revisiter les approches mécanistico-statistiques (typiquement appliquées à des phénomènes et données d'écologie et d'épidémiologie) via (i) le développement de méthodes de résolution et d'estimation simultanées d'EDP et modèles intégraux par des techniques d'apprentissage profond, ainsi que (ii) l'intégration dans l'approche mécanistico-statistique des données massives informant les dynamiques écologiques et épidémiologiques à de grandes échelles spatiales en développant pour ce cadre de modélisation d'inspiration mécaniste les techniques de régularisation, d'auto-encodage et de méta-modélisation. Parallèlement à ces pistes de recherche mobilisant les

démarches d'apprentissage, nous envisageons de renforcer l'assise théorique de la modélisation mécanistico-statistique avec notamment (iii) la mise au point d'un cadre mathématique rigoureux pour l'estimation de paramètres de modèles d'EDP, s'appuyant sur des techniques issues des problèmes inverses, des probabilités et des statistiques théoriques,

- Il s'agit également de développer une vision spatio-temporelle et multi-échelle des réseaux trophiques et fonctionnels, des microbiomes, des pathobiomes ou encore des phytobiomes via des processus ponctuels spatio-temporels ou des réseaux probabilistes spatio-temporels définis sur des espaces de projection. Ces espaces de projection devront refléter l'essentiel des informations pertinentes contenues dans les données brutes de diversité caractérisant les communautés susmentionnées. Ils pourront être obtenus en adaptant et développant des techniques de réduction de la dimension issues du domaine de l'apprentissage automatique. A cet égard, des travaux ont été initiés récemment dans l'unité sur les auto-encodeurs appliqués à des données de diversité génomique de pathogènes afin de reconstruire les transmissions de maladie. Au-delà de ce cas d'étude, le développement de la vision susmentionnée permettra de représenter, caractériser et nowcaster à plusieurs échelles les dynamiques spatio-temporelles des communautés (réseaux trophiques, microbiomes...), et de les prédire sous divers scénarios de contrôle avec comme cadres d'application principaux les émergences de maladies, les dynamiques de pathogènes multi-hôtes ou encore les processus de conservation.
- Il s'agit enfin, sur le plan informatique et via la mobilisation d'une partie des ingénieurs de l'unité, d'explorer les architectures de calcul autres que CPU à la base de notre cluster de calcul, pour étudier les gains associés au calcul GPU notamment (calcul de produits de convolution dans des équations non locales, MCMC, apprentissage profond appliqué à nos sujets). L'unité a ainsi investi en 2022 dans un serveur de calcul GPU pour mener des premiers tests.

Comme nous l'avons mentionné ci-dessus, depuis plusieurs années, l'unité fait remonter aux tutelles deux demandes de postes : un poste de CR "Méthodes d'apprentissage en statistiques spatiales", et un IE spécialisé en calcul scientifique. Les points évoqués ci-dessus alimenteront les prochaines demandes de recrutement, principalement de CR et de doctorants, avec des compétences en modélisation spatio-temporelle (EDP, processus et réseaux stochastiques), analyse mathématique, statistique inférentielle, apprentissage automatique...

Référence 2. La production scientifique est proportionnée au potentiel de recherche de l'équipe et répartie entre ses personnels.

Avec en moyenne 13.8 chercheurs ou chargés de mission sur la période évaluée et 206 articles sur la période, le taux de publication des chercheurs de l'équipe Recherche est de 15 articles par chercheur, soit 2.5 articles par chercheur et par an. Parmi ces travaux, 40% correspondent à des travaux en mathématiques, 41% à des travaux interdisciplinaires et 19% à des travaux en biologie (généralement dans le cadre d'une collaboration portée par un collègue biologiste).

Le taux de publication varie entre 0.8 et 6.5 publications/an/chercheur, avec une moyenne de 2.7 (> aux 2.5 mentionnés ci-dessus car dans ce cas certaines publications sont comptées plusieurs fois), tous les chercheurs participent donc à la production de l'unité. Sur les 206 publications, 89 sont co-signées par un doctorant, et 29 par un post-doc. Comme nous l'avons déjà mentionné en Dom 2 Ref 2, les étudiants ayant effectué leur thèse à BioSP (et qui étaient hébergés à BioSP) pendant la période évaluée ont publié 25 articles, pour un total d'environ 30 doctorant*année, soit environ 1 article par étudiant et par an. Sur les 10 soutenances de thèses durant la période, 6 étudiants avaient démarré leur thèse après 2016. Leur travail s'est donc entièrement déroulé au cours de la période évaluée ; ils ont en moyenne publié 2.8 articles pendant leur thèse (entre 1 et 4).

Les scientifiques de l'équipe Recherche ont en outre participé à la rédaction de 16 chapitres d'ouvrages, au développement de 24 logiciels (notamment des packages R, des webapps R-Shiny) et de 4 bases de données.

Comme nous l'avons expliqué dans la partie *Prise en compte des recommandations du précédent rapport* au début de ce document, l'équipe Recherche cherche volontairement à distribuer l'ingénierie équitablement dans les différents axes disciplinaires et priorités scientifiques, afin d'éviter une déconnection aux activités de recherche de l'unité. Ainsi, les ingénieurs de l'équipe contribuent soit directement à la production d'articles (11 articles co-signés sur la période) soit indirectement en mettant à disposition des outils de calcul (cluster, cf Dom 2 Ref 4) ou de collaboration (e.g., GitLab, JupyterHub, hébergement Apps R-Shiny), ou des webapps associées aux articles [LOG-RRP21,PUB-RPR18,LOG-BRL21,PUB-WLR19,LOG-RK19, PUB-OMG18]. Parallèlement à cela, les ingénieurs de l'équipe mènent des projets d'ingénierie fédérateurs, comme le projet SK8 visant une infrastructure matérielle et logicielle de développement et d'hébergement d'applications Web R-Shiny ou le projet MSE détaillé ci-dessus.

Les scientifiques de BioSP sont régulièrement invités à présenter leurs travaux aux autres collègues de l'unité au moment du Café des Sciences hebdomadaire, au cours des visites annuelles des départements ou de la réunion hors-les-murs annuelle, via les événements organisés dans le cadre de la FRUMAM et des réseaux scientifiques portés par l'unité (cf. Dom 2 Ref 1). Ces événements sont suivis de discussions informelles permettant aux scientifiques de BioSP d'améliorer leur connaissance de l'état de l'art et de garantir une bonne qualité des travaux.

Equipe n° 2 (OPE)

Référence 1. La production scientifique de l'équipe satisfait à des critères de qualité.

Les productions de l'équipe OPE découlent des besoins des partenaires de la Plateforme ESV et de ses interactions avec la recherche (cette deuxième dimension est peu développée dans ce rapport compte-tenu de la jeunesse de l'équipe et du temps long de la recherche).

Ces productions sont diverses et sont classées en trois catégories :

- Des productions écrites, en particulier des fiches de reconnaissances d'organismes nuisibles [COL-OPE21a] qui sont une ressource officielle et harmonisée des méthodes de reconnaissance, de piégeage et de prélèvements de ces organismes ; des bulletins de veille sanitaire et scientifique internationale [COL-OPE21b] produits sur une base régulière ; d'un rapport co-écrit par notre équipe et un collectif d'experts concernant la mission Occitanie pour *Xylella fastidiosa* (élément non diffusable pour le moment) ; d'un guide qualité des données de surveillance rédigé dans le cadre d'un groupe de travail de la plateforme.
- Des applications web à destination du public (grand public, décideurs publics et privés, laboratoires d'analyse, groupes de travail de la plateforme, partenaires de la plateforme) avec ou sans accès restreint. Ces applications servent à la visualisation de données et d'analyses, concernant p ex la surveillance du nématode du pin [CAR-OPE19] (une version enrichie en accès restreint est également disponible pour des partenaires publics et privés ciblés [CAR-OPE21b]), du virus émergent ToBRFV (*Tomato brown rugose fruit virus*) [CAR-OPE20a] (**Fig. 10**) et de la processionnaire du pin [CAR-OPE21a], ou encore la cartographie des zones à risque d'entrée et d'introduction du nématode du pin en France [CAR-OPE20b]. Le guide pratique sur la qualité des données de surveillance auquel l'équipe OPE a largement participé, dans la création du contenu, a été rendu public en ligne sous la forme d'un wiki.
- Des applications web internes permettent de faciliter les activités de la Plateforme ESV. C'est le cas de l'application de Data-management pour la veille sanitaire et scientifique internationale, qui dispose d'une interface graphique facilitant le travail des "veilleurs" (tri et filtre des articles). C'est également le cas de l'application Fiche-diagnostique permettant de centraliser des connaissances/expertises, d'automatiser la production de fiches de reconnaissance, et de les mettre à disposition du grand public et des techniciens sur le terrain.

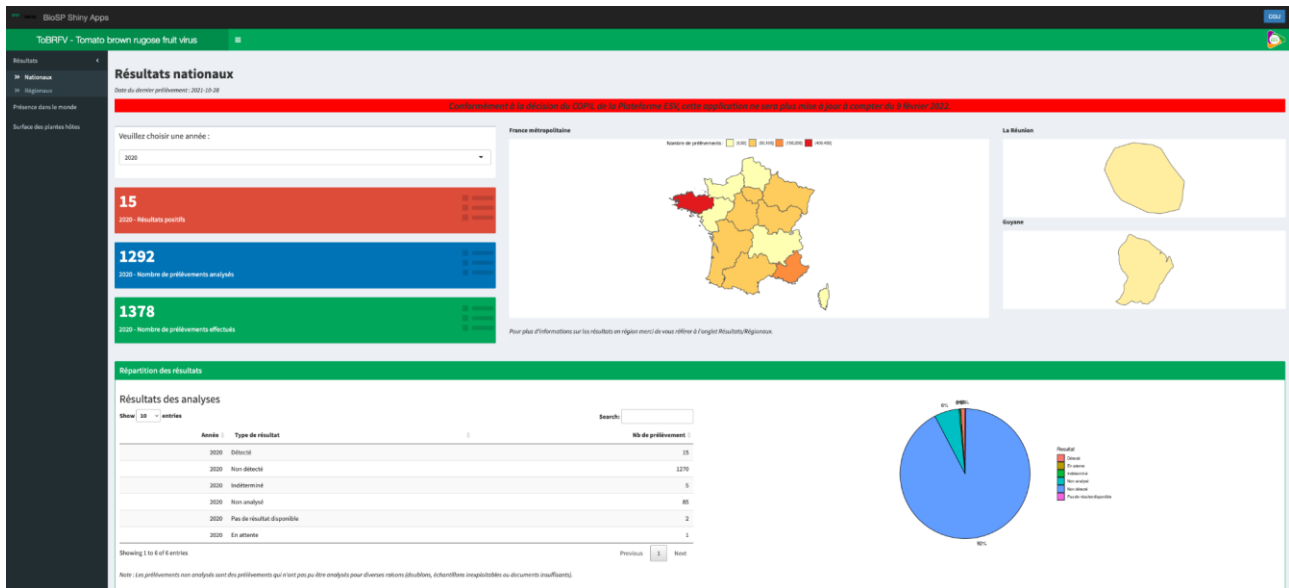


Fig. 10. Tableau de bord et application de visualisation des données de surveillance du Tomato brown rugose fruit virus (ToBRFV) [CAR-OPE20a] à l'échelle nationale et régionale.

L'ensemble des productions et applications publiques sont disponibles sur le site web de la Plateforme ESV (<https://plateforme-esv.fr/>) hébergé sur un data center d'INRAE, développé et maintenu par l'équipe OPE en collaboration avec les autres membres (Anses, Cirad) de l'équipe opérationnelle de la Plateforme ESV. La Plateforme ESV valorise également ces productions sur le réseau social Linked'in, géré par la chargée de communication de l'équipe.

L'ensemble de ces productions apportent de nouvelles connaissances dans la sphère professionnelle de la santé du végétal, contribuent à améliorer la surveillance des organismes nuisibles réglementés et servent de base pour des actions de communication autour de ces thématiques.

Référence 2. La production scientifique est proportionnée au potentiel de recherche de l'équipe et répartie entre ses personnels.

L'ensemble des compétences de l'équipe OPE (épidémiologie, informatique, bio-statistique, communication) sont mobilisées pour la plupart des productions développées, mobilisation permise par l'investissement soutenu de chacun des agents de l'équipe.

Depuis la création de la plateforme en 2018 (et eu égard au recrutement progressif des 7 agents de l'équipe étalé entre fin 2018 et 2021), 31 fiches de reconnaissance ont été publiées [COL-OPE21a] sur le site web et 25 se trouvent dans un état proche de leur publication à la fin de l'année 2021 et, dans le périmètre de la veille sanitaire internationale, 96 bulletins hebdomadaires, 29 bulletins mensuels et 1 « Point sur » (i.e., une note de synthèse sur un sujet d'actualité particulier) ont été produits (disponible sur le site web et via une liste de diffusion) [COL-OPE21b]. La communication autour de la plateforme s'articule selon quatre thèmes : la santé du végétal, l'identité de la plateforme, les expertises de la plateforme et la vie et mobilité de la plateforme, 51 articles ont été écrits et relayés sur le réseau social Linked'in.

En terme de développement d'outils à destination des divers publics, l'équipe comptabilise sur la période le développement de 3 applications de visualisation "Shiny" (deux pour le nématode du pin et une pour le ToBRFV), de 3 dashboards (nématode du pin, processionnaire du pin, *Bretziella fagacearum*) et d'un wiki (format web du guide qualité des données). 2 autres applications pour la veille sanitaire et pour la construction des fiches diagnostiques ont été développées dans l'objectif de faciliter le travail de l'équipe.

Unité BioSP (les 2 équipes)

Référence 3. La production scientifique de l'unité respecte les principes de l'intégrité scientifique, de l'éthique et de la science ouverte.

Comme mentionné ci-dessus, les scientifiques de BioSP sont régulièrement invités à présenter leurs travaux aux autres collègues de l'unité, ce qui constitue un gage important de qualité et d'intégrité. Concernant les publications, le choix du support est du ressort des auteurs et la direction d'unité ne peut qu'orienter ce choix, par exemple en relayant l'initiative autour des revues PCI (Peer Community In, cf ci-dessous). Les “pauses café” et les repas du midi sont souvent l'occasion de discussion nourries entre chercheurs, concernant la qualité des revues, le moment où l'on peut considérer qu'un article est prêt à être publié, la longueur des articles (articles denses avec de nombreux résultats vs articles présentant un résultat marquant), l'importance de faire un état de l'art exhaustif, l'ordre alphabétique... ces discussions permettent de bien informer les chercheurs juniors quant à la qualité des supports et la rigueur nécessaire dans la présentation des résultats, mais montrent qu'il est difficile de se baser sur une règle simple, qui devra dépendre de la discipline (collaboration avec des biologistes ou non), du moment de la carrière (permanent ou non-permanent), du type de résultat (avec potentiellement une large audience ou non), de l'existence d'un papier compagnon plus technique, de l'urgence à publier (comme dans le cas de nos résultats obtenus pendant l'épidémie de Covid-19), du souhait d'être bien cité... L'unité ne finance pas les déplacements pour assister à des conférences dites “prédatrices” ; nous n'avons de fait jamais eu à faire face à ce type de demande. En cohérence avec le 2e Plan national pour la science ouverte, l'unité vise à :

- Généraliser l'accès ouvert aux publications, avec 70% des publications disponibles en “Open access”. Il s'agit majoritairement d'articles déposés sur des serveurs de preprint (HAL, ArXiv, bioRxiv, medRxiv, ...). L'unité donne également la possibilité de financer la publication dans des revues open access (p ex *Royal Society Open Science*), des revues de sociétés savantes (p ex *Genetics*) ou de payer l'option open access dans les revues qui le permettent. La direction d'unité relaye régulièrement l'initiative PCI, lancée par des collègues INRAE et consistant en un service de recommandation de préprints basé sur des évaluations par les pairs. Sur la période, 3 preprints de l'unité ont été recommandés par PCI [PUB-LMA20,PUB-PFA21,PUB-HML21] ;
- Ouvrir et promouvoir les codes sources produits par la recherche. Ces dernières années, en cohérence avec les demandes des revues, nos codes sont généralement partagés via des répertoires de dépôt tels que Zenodo, Open Science Framework ou des dépôts GitLab publics ;
- Partager le plus souvent possible les données associées aux articles, sous forme de matériel supplémentaire accompagnant les articles et/ou dans des outils d'archivage tels que Zenodo, Open Science Framework et le dataverse INRAE.
- Transformer plus globalement les pratiques pour faire de la science ouverte le principe par défaut, suivant le modèle “FAIR” (Facile à trouver, Accessible, Interopérable, Réutilisable) via, d'un point de vue général, l'utilisation de technologies pour faciliter le partage et la reproductibilité, et en particulier l'ouverture publique des dépôts de code, la construction de pages web vitrines, l'intégration continue pour le packaging et les tests de régression, la diffusion des bonnes pratiques de codage.

En lien avec les points susmentionnés, deux ingénieurs de l'unité ont été nommés “référénts données opérationnels” fin 2021 (S Nicoux et E Walker). Ils sont notamment chargés de développer et mettre en œuvre une culture commune des bonnes pratiques en matière de gestion des données.

Par ailleurs, l'équipe OPE à travers son Système d'Information a mis en place des mesures de sécurisation afin de respecter l'intégrité des données collectées et des travaux menés. Elle travaille étroitement avec la RSSI INRAE et le responsable CNIL afin de respecter le RGPD des différentes données traitées. L'équipe OPE participe également au principe de science ouverte dans le cadre de ses activités de mise à disposition des données de surveillance à des fins de recherche.

Synthèse de l'autoévaluation, Domaine 3

Forces. L'unité définit de façon dynamique des priorités scientifiques qui entrent en résonance avec les grands objectifs scientifiques d'INRAE et des départements MathNum et SPE, et qui permettent d'exploiter au mieux les compétences disciplinaires des agents. Tirés par les applications issues de collaborations avec des

biologistes, les travaux de l'équipe Recherche ne sont pas figés dans l'analyse d'un nombre réduit de questions théoriques. Ceci dit, malgré une part importante de travaux interdisciplinaires et de forces centrifuges exercées par les collaborations avec les autres unités INRAE et hors INRAE, l'équipe reste solide sur ses bases théoriques avec des travaux en statistiques, probabilités et analyse des EDP paraissant dans des revues reconnues et en collaboration avec des collègues issus du monde académique ayant un très fort rayonnement international (comme H Berestycki, F Hamel, E Pardoux ou H Rue, E Porcu, J Moller). Plusieurs de nos collaborateurs biologistes ont également une vaste aura internationale (comme L Belzunces, V Bretagnolle, C Morris, A-L Laine, O Martin, O Ronce, A Lang, P Thrall ou encore C. Stefanescu). Ainsi, l'équipe présente une production scientifique équilibrée entre travaux en mathématiques / interdisciplinaires / en biologie, bien répartie entre chercheurs, s'appuyant significativement sur les compétences en ingénierie présentes dans l'unité et valorisant nos collaborations internationales. Les étudiants en doctorat et post-doctorat s'y forment généralement un solide dossier de publication.

L'existence d'une unique équipe Recherche, comportant 3 axes disciplinaires a priori éloignés, et la structuration croisée de BioSP entre priorités scientifiques et axes disciplinaires sont favorables à l'émergence de travaux originaux, couplant des méthodes issues de fronts de sciences dans chacune des disciplines. On peut citer les travaux autour des approches mécanistico-statistiques, qui font appel aux compétences en systèmes dynamiques, en statistiques, en simulation numérique et souvent en écologie-épidémiologie. De même, les travaux récents autour de la modélisation de champs spatio-temporels par EDPS jettent un nouveau pont entre les statistiques spatiales et les systèmes dynamiques (ici EDP avec Laplacien fractionnaire et équations intégro-différentielles) en revisitant la construction EDPS des champs spatiaux sous un angle mécanistique [PUB-RAS21]. On compte ainsi 17 publications dans lesquelles statisticiens et EDPistes sont impliqués (e.g. [PUB-GCC17, PUB-RBB20, PUB-RDB21]). Le recrutement récent de R Forien et les éléments de projet autour des approches mécanistico-statistiques témoignent de la volonté de l'équipe d'accroître encore la perméabilité entre ses différents axes disciplinaires.

L'équipe OPE est positionnée sur des enjeux sociétaux importants, notamment en ce qui concerne les organismes de quarantaine, et peut s'appuyer sur une équipe multidisciplinaire stabilisée, les derniers recrutements de permanents ayant eu lieu en 2021. Son implication dans des projets de recherche (projets communs avec l'équipe Recherche de BioSP et projets avec des équipes externes, par ex., le projet CoACT2 avec les UMR BFP et SAVE et le projet TIERS-ESV avec l'UR MaIAGE) permet une émulation scientifique au sein de l'unité BioSP et d'INRAE autour des questions d'épidémiologie.

Faiblesses. L'équipe Recherche n'a pas publié dans les revues généralistes les plus reconnues comptant dans le classement de Shangai (Nature, Science). Le partage de l'information autour du choix des supports de publication se fait de façon informelle. Ce travail d'auto-évaluation nous amènera sans doute à proposer des sessions d'information destinées aux nouveaux arrivants sur le thème des revues prédatrices de l'open access, des initiatives de type PCI...

L'équipe OPE n'a pas encore de participation significative dans des publications scientifiques mais son implication récente (à partir de 2020) dans des projets de recherche le permettra à court terme. Ses productions ne sont pour l'heure pas citées de manière significative mais nous travaillons à leur DOIisation et à l'apposition de QR-codes sur certaines d'entre elles (jeux de données, fiches de reconnaissance, rapports). Par ailleurs, la Plateforme ESV (à laquelle contribue l'équipe OPE) a deux analogues : la plateforme de surveillance de la chaîne alimentaire (créée en 2018 en même temps que la Plateforme ESV) et la plateforme d'épidémiologie en santé animale (créée en 2011 et renouvelée en 2018). Si l'existence de ces deux plateformes est bénéfique pour des partages d'expérience et la mise en œuvre de travaux communs ponctuels (dans le cadre de groupes de travail inter-plateformes et du CATI IMOTEP notamment), l'incitation actuelle (au niveau du MAA) à une mutualisation des systèmes d'information des trois plateformes peut être considérée comme un risque. Une telle mutualisation limiterait vraisemblablement l'agilité de l'équipe OPE dans l'exercice de ses missions. L'identification de travaux communs ponctuels répondant aux priorités de tous, et le transfert (d'une plateforme à l'autre) de technologies et d'outils lorsque cela s'avère pertinent semblent préférables à une mutualisation.

Domaine d'évaluation 4 : Inscription des activités de recherche dans la société

Comme nous l'avons rappelé en préambule du Dom 3, l'équipe Recherche contribue essentiellement aux activités du Dom 3, et l'équipe OPE à celles du Dom 4. Les deux équipes interagissent régulièrement, en collaboration avec des partenaires académiques dans des projets de recherche (typiquement avec des unités d'INRAE et du Cirad) et des partenaires non académiques tels que des organisations nationales et internationales (MAA, Anses, EFSA, FAO). Globalement, l'unité travaille avec des partenaires divers et ses membres se prêtent volontiers aux activités de vulgarisation auprès du grand public.

Equipe n° 1 (Recherche)

Référence 1. L'équipe se distingue par la qualité de ses interactions non-académiques.

En collaboration avec l'équipe Géostatistique à MINES Paris, BioSP est porteur d'une chaire de mécénat Geolearning portant sur les méthodes et outils pour la géostatistique, les événements extrêmes et le machine learning au service de la transition écologique. Cette chaire sera financée par 4 entreprises mécènes (ANDRA, BNP-Paribas, CCR et SCOR) pour un montant total de 1.8 M€ de 2022 à 2027. Le montage de cette chaire (trouver les entreprises mécènes, identifier les axes de recherches et construire les premiers projets) a démarré en 2020 et a donc nécessité deux ans de travail.

BioSP mène également des travaux en statistiques appliquées pour développer des méthodes adaptées aux données de sciences participatives. Leur analyse est délicate du fait d'un effort d'échantillonnage inconnu et hétérogène dans l'espace et le temps. Ceci s'est traduit par le coencadrement de la thèse de Christophe Botella (valorisation des données issues de l'application Pl@ntNet) et l'encadrement de celle de Florian Lasgorceux (en partenariat avec le parc national des Ecrins). De plus, BioSP porte le réseau CiSStats depuis 2011. Ce réseau a pour objectif de fédérer les chercheurs travaillant sur l'analyse de données issues des sciences participatives, mais aussi de transférer ces méthodes vers les porteurs d'enjeux non académiques et de construire avec eux des questionnements de recherche pertinents. Enfin, depuis peu, la priorité thématique #Obs-model de BioSP porte spécifiquement sur l'analyse de telles données.

Le développement des recherches sur l'activité des abeilles s'articule sur un partenariat étroit entre chercheurs (ce thème est porté par A Kretzschmar, qui a actuellement un statut de chargé de mission), associations de développement et professionnels de l'apiculture (11 apiculteurs, ADAPI en France, UNAAPi en Italie). L'objectif de cette collaboration est de construire des outils utiles et efficaces pour aider à la gestion durable des exploitations apicoles. Pour permettre aux professionnels de l'apiculture d'améliorer le développement de leur stratégie d'exploitation, des outils de cartographie et d'analyse spatio-temporelle sont proposés à partir d'instrument connectés (balances), de mise à disposition de base de données et de diffusion d'information sur [site Web en français](#) et en [italien](#). Le développement des outils expérimentaux et scientifiques ont permis de contribuer au débat général sur le déclin des abeilles en y apportant un point de vue objectif à long terme : développement des observatoires des miellées de lavandes et tournesol (respectivement 14 et 7 années). Dans le cas de l'observatoire de la miellée de lavandes, l'observatoire devient, pour les professionnels de l'apiculture, un outil pérenne. Notons enfin que l'équipe participe régulièrement depuis 15 ans aux journées techniques des Associations régionales de développement de l'Apiculture et propose une assistance à la mise en place des expérimentations et à la réalisation de l'analyse des données.

Concernant la santé des plantes et l'épidémiologie, l'équipe Recherche est impliquée dans plusieurs partenariats, à travers la Plateforme ESV, mais pas seulement. Ainsi, des chercheurs de BioSP sont impliqués dans plusieurs groupes de travail de la plateforme (groupes mis en place par la Plateforme ESV sur des thématiques biologiques telles que la surveillance de *Xylella fastidiosa* et des travaux méthodologiques tels que la qualité des données) et apportent par ce biais leur expertise scientifique au Ministère de l'Agriculture et de l'Alimentation, aux chambres d'agriculture, aux instituts techniques ou encore à l'EFSA [EXP-S18,EXP-M19,EXP-M21,EXP-S21]. Cette expertise est notamment développée dans plusieurs projets incluant pour la plupart des partenaires non-académiques (projet UE XF-ACTORS, projet UE IPM-Popillia, projet MAA-INRAE sur *Xylella fastidiosa*). Au-delà du cadre européen, une collaboration a été initiée en 2021 avec la FAO sur la surveillance de *Spodoptera frugiperda* en Afrique, l'identification des facteurs de risque et des capacités de dispersion.

L'équipe a également été impliquée sur l'étude des effets non intentionnels des pratiques agricoles en collaboration avec des partenaires non académiques, d'abord dans le cadre de la biovigilance à travers le pilotage par P Monestiez du groupe de travail national STATS 500-ENI (Plan ECOPHYTO), ensuite à travers un projet financé par l'EFSA (BriskaR-NTL) sur les effets des toxines contenues dans le pollen de maïs génétiquement modifié sur des espèces non cibles (qui ne sont pas nuisibles à la culture, mais qui sont impactées par les toxines). Typiquement, le projet BriskaR-NTL a permis de produire un rapport [PRO-BLS21] et un package R (briskaR : [LOG-BRL21]) à destination des gestionnaires des risques, en plus d'un article de recherche [PUB-BWL2021].

Sur la période, l'équipe Recherche a accueilli 4 doctorants financés à 50% par la région PACA qui requiert un partenariat avec des acteurs socio-économiques de la région sur la durée de la thèse. Par cet intermédiaire, l'équipe Recherche a interagi avec l'AFIDOL, la chambre d'agriculture des Alpes Maritimes, CTIFL, ITAB, IAS (entreprise de production d'auxiliaires de lutte biologique). Ce cadre de partenariat a principalement permis des partages d'expérience.

Référence 2 : L'équipe développe des produits à destination du monde socio-économique.

Sur la thématique "abeilles", un important cahier technique a été réalisé dans le cadre du projet Innov'API [PRO-K21]. La validation d'une technique nouvelle pour gérer un des problèmes sanitaires majeurs de l'apiculture (Varroa) offre une solution alternative durable aux professionnels pour pallier le développement de résistance des médicaments vétérinaires de synthèse. La diffusion des résultats est partie prenante du projet, avec la mise à disposition de tous les résultats au fil du projet par la mise à jour d'un [site Web dédié](#).

L'équipe a produit sur la période plusieurs rapports d'expertise scientifique, notamment pour des partenaires publics nationaux ou internationaux tels que l'ANSES, l'EFSA, l'EPPO, la FAO ou le Ministère de l'Agriculture et de l'Alimentation. Il s'agissait notamment d'utiliser la modélisation statistique [EXP-M19, EXP-M21] ou par EDP [EXP-BR19] pour établir des cartes de risque en épidémiologie ou concernant des invasions biologiques.

Référence 3. L'équipe partage ses connaissances avec le grand public et intervient dans des débats de société.

Les membres de l'équipe ont écrit des articles de vulgarisation scientifique, dans le cadre de la fête de la science ou du Comité international des jeux mathématiques (CIJM) [MED-K16, MED-R20]. Au cours de l'épidémie de Covid-19, et dès le début de l'année 2020, l'équipe a proposé une analyse régulière des données épidémiques sur un [blog](#). Les travaux de l'équipe sur ce thème ont donné lieu à [2 communiqués de presse INRAE](#), un [communiqué de presse de l'EHESS](#), et ont été intégrés à un dossier de presse digital regroupant les recherches INRAE sur Covid-19 (<http://covid19.inrae.fr/dossierdepresse/>).

Dans le cadre d'une collaboration de longue date entre BioSP, l'unité de Pathologie Végétale d'INRAE Avignon et le CSIRO de Canberra, cette période a été marquée par le développement d'un logiciel (sous forme du package R "landsepi", [LOG-RRP21]) associé à une interface web permettant la simulation d'épidémies à l'échelle d'un territoire agricole. Ce logiciel est au cœur de plusieurs projets internes à INRAE mais aussi ANR et Ecophyto, il a été retenu également par le département SPE comme un fait marquant pour la rédaction du document "2010-2019: 10 ans de recherche au département Santé des Plantes et Environnement". Le package "landsepi" via son application web est régulièrement utilisé comme support de cours et de TD dans des masters nationaux et internationaux ou des écoles d'ingénieurs. De plus, il a été présenté plusieurs fois lors des "fêtes de la science" afin de sensibiliser le grand public à l'intérêt de la diversification des paysages agricoles.

Sur la thématique "abeilles", notons l'organisation d'un Webinaire en 2021 qui a réuni plus de 1000 participants, dont de nombreux professionnels de l'apiculture à la fin du projet Innov'API.

Les recherches de BioSP sur la reconstruction des transmissions à partir de données génomiques ont par ailleurs été présentées à un large public au salon de l'agriculture (SIA Paris) en 2017, dans le cadre du colloque de l'INRA traitant cette année-là du sujet One Health [MED-S17]. Une intervention devant le même type de public a été réalisée en 2018 lors du salon Med'Agri à Avignon, dans une session publique intitulée "Le numérique au service de l'agriculture : concepts et réalités" [MED-S18]. Deux scientifiques de BioSP ont

également contribué à un dossier de presse sur One Health en 2020 [MED-PS20] et l'un des deux sujets de BioSP a été présenté lors d'une webconférence à destination des journalistes [MED-S20].

Enfin, on notera que l'unité accueille régulièrement des collégiens et lycéens, en moyenne 3 à 4 élèves /an. Sur la période évaluée, 20 collégiens de 3ème et 5 lycéens ont été accueillis.

Synthèse de l'autoévaluation, Domaine 4, équipe Recherche

Forces. L'équipe Recherche entretient des collaborations diversifiées avec des acteurs du monde non académique appartenant au monde de la ré-assurance (BNP-Paribas, CCR et SCOR) via la chaire Geolearning, les gestionnaires du Parc national des Ecrins, les professionnels de l'apiculture et des organisations nationales et internationales intervenant dans le secteur de la santé des plantes. Les interactions entre équipe Recherche et équipe OPE ont renforcé et pérennisé les collaborations avec les acteurs de la sécurité alimentaire et de la santé végétale (Ministère de l'Agriculture et de l'Alimentation, Anses, EFSA, FAO). Au travers de la priorité scientifique #Obs-model et du réseau CiSStats, l'équipe contribue au lien entre recherche en statistiques et porteurs d'enjeux non académiques autour des données de sciences participatives. Enfin, depuis peu, les membres de l'équipe contribuent à l'interaction science-société en publiant des articles de vulgarisation, en utilisant des supports accessibles (blogs, interface web adossée à un logiciel, communiqués de presse) et en participant à des animations à destination des scolaires (fête de la science, accueil d'élèves du secondaire).

Faiblesses. Si des collaborations ponctuelles existent entre des chercheurs de l'équipe et des spécialistes des abeilles, la pérennité des partenariats non-académiques autour de la thématique "abeilles" repose sur un chargé de mission qui quittera vraisemblablement l'unité dans les prochaines années.

Equipe n° 2 (OPE)

Référence 1. L'équipe se distingue par la qualité de ses interactions non-académiques.

Le périmètre potentiel des travaux de la Plateforme ESV est l'ensemble des dangers sanitaires pouvant impacter les végétaux ainsi que les effets non intentionnels des pratiques agricoles. Ceci dit, sur la période évaluée, le MAA et le comité de pilotage de la Plateforme ESV ont priorisé la surveillance d'organismes nuisibles phares ou à fort risque d'émergence. Ainsi, l'équipe OPE travaille spécifiquement sur l'amélioration de la surveillance des organismes nuisibles et déperissements suivants : *Xylella fastidiosa*, nématode du pin, déperissement de la vigne, ToBRFV, processionnaire du pin, HuangLongBing, FocTR4 (*Fusarium oxysporum f.sp. cubense* race 4), *Bretziella fagacearum*.

Les activités de l'équipe OPE peuvent être décrites selon les 5 missions suivantes, en cohérence avec la convention cadre signée en 2018 pour une durée de 10 ans :

1. *Expertise scientifique et phytosanitaire* avec la présence de 2 épidémiologistes au sein de l'équipe qui apportent leurs connaissances dans les différents cas d'études sur des organismes nuisibles spécifiques et à travers la veille sanitaire internationale.
2. *Gestion du système d'information* avec la mise en place d'un SI sécurisé mettant à disposition des applications web, un site web, une base de données où sont centralisées les données de surveillance et l'ensemble de l'écosystème des données de la plateforme.
3. *Traitement et analyse des données* permettant de connaître l'état sanitaire du territoire, de disposer de cartographies des dispositifs de surveillance en place et de l'évolution des épidémies, ainsi que de comprendre les mécanismes associés à ces organismes nuisibles (dispersion temporelle et spatiale).
4. *Communications scientifiques et diffusion d'informations* via les bulletins de la VSI, les actualités publiées sur le site web et le Linked'in de la Plateforme ESV, les fiches de reconnaissance (un exemple de fiche, celle de *Spodoptera frugiperda* est disponible dans le portfolio), ou encore des interventions lors d'événements grand public (salon de l'agriculture à Paris en 2020, salon Med'agri à Avignon en 2020), lors de séminaires professionnels ou lors de conférences scientifiques. Par ailleurs, dans ce cadre de diffusion, une activité de mise à disposition des données de surveillance à des fins de

- recherche a été développée par la Plateforme ESV, les données étant disponible sur le portail data INRAE (un exemple de jeu de données pour *Xylella fastidiosa* est disponible dans le portfolio).
5. *Évaluation des dispositifs de surveillance* afin de les améliorer, par exemple en comparant les zones surveillées actuellement avec les zones à risque d'entrée ou d'introduction (entrée + établissement) des organismes nuisibles.

Dans le cadre de l'accomplissement de ces cinq missions, l'équipe OPE entretient un lien étroit avec les 6 autres partenaires de la Plateforme ESV : MAA, Anses, APCA, FREDON France, ACTA - les instituts techniques, Cirad, dont 5 sont des partenaires non académiques. Le lien avec l'Anses est particulièrement fort puisque l'équipe opérationnelle de la Plateforme ESV est composée de 8,5 ETP dont les 7 agents INRAE (équipe OPE), 3 agents Anses et 1 agent Cirad (pour les 1,5 ETP restants).

Par ailleurs, la Plateforme ESV a mis en place des groupes de travail sur des thématiques biologiques telles que la surveillance de *Xylella fastidiosa* et des travaux méthodologiques tels que la qualité des données (une quinzaine de groupes de travail sont actuellement actifs). Au sein de chacun de ces groupes, la majeure partie des membres est issue d'organisations non académiques (organismes de prospections, instituts techniques agricoles, professionnels des filières agricoles) avec lesquels l'équipe OPE est amenée à interagir. Ces acteurs sont en particulier les destinataires principaux (avec le MAA et les services de l'État en régions et dans les DROM - SRAL/SALIM) de nos outils résultant des missions de traitement de données et d'évaluation et amélioration des dispositifs de surveillance.

Référence 2 : L'équipe développe des produits à destination du monde socio-économique.

L'équipe OPE dont l'activité est pleinement dédiée à la Plateforme ESV, vient en ce sens en appui aux politiques publiques (ministère et services en région métropole/DROM), aux acteurs de la santé végétale (organismes de prospections, instituts techniques agricoles, professionnels des filières agricoles).

Cet appui se concrétise en proposant des outils méthodologiques, des outils de visualisation cartographique et des outils de veille pour détecter le plus tôt possible l'arrivée des maladies sur le territoire, pour délimiter les zones à prospecter et améliorer les dispositifs de surveillance, pour aider à la reconnaissance et communiquer autour des organismes nuisibles. Tous ces outils intéressent les organismes de prospection et les professionnels des filières ainsi que le ministère de l'agriculture et les services en région.

Référence 3. L'équipe partage ses connaissances avec le grand public et intervient dans des débats de société.

L'activité de communication de la plateforme et donc de l'équipe OPE est tournée à la fois vers les politiques publiques, les acteurs de la santé végétale et le grand public. La présence de l'équipe au Salon Internationale de l'Agriculture en 2020 a été l'un des premiers tremplins de communication pour la Plateforme ESV auprès du grand public. Certaines des productions de l'équipe OPE sont accessibles au grand public (fiches de connaissance, avancées sur les travaux des groupes de travail, cartographie de la surveillance d'organismes nuisibles, actualités) via le [site web](#) de la plateforme et le Linked'in permettant de communiquer très largement autour de la santé des végétaux et de sensibiliser le plus grand nombre à cette thématique.

Synthèse de l'autoévaluation, Domaine 4, équipe OPE

Forces. L'équipe OPE interagit de manière régulière avec des agents d'autres organisations via des groupes de travail (MAA, services de l'Etat en région, Cirad, Anses, chambres d'agriculture, instituts techniques...) et est impliquée dans de multiples co-productions. Le format de la plateforme place l'équipe au cœur d'un réseau rassemblant de multiples acteurs de la santé du végétal et sa visibilité s'étend de la sphère ministérielle au grand public.

Faiblesses. De par la multitude des besoins et des publics, de nombreux fronts sont à mener en parallèle. Cela demande une adaptation perpétuelle de l'organisation du travail de chacun et une agilité élevée de l'équipe.

APPENDICES

Liste des acronymes utilisés dans le document

ACCAF	Adaptation aux changements climatiques de l'agriculture et des forêts
AMU	Aix-Marseille Université
CT	Champs thématiques des départements
DSI	Direction Systèmes d'Information
ESV	Epidémiosurveillance en Santé Végétale
FRUMAM	Fédération de Recherche des Unités de Mathématiques de Marseille
GOS	Grands objectifs scientifiques des départements
GU	Gestionnaire d'Unité
MAA	Ministère de l'Agriculture et de l'Alimentation
MathNum	Département mathématiques, informatique, sciences de la donnée et technologies du numérique
MP	Métaprogramme INRAE
OPE	Équipe opérationnelle de la Plateforme ESV
OS	Orientations scientifiques INRAE
PCI	Peer Community In
PS	Priorité Scientifique de l'unité BioSP
RSSI	Responsable de la sécurité des systèmes d'information
SALIM	Service de l'Alimentation
SRAL	Service Régional de l'Alimentation
SPE	Département Santé des Plantes et Environnement

PORTFOLIO

Le portfolio de l'équipe Recherche se compose de 7 éléments. Nous proposons 5 publications correspondant aux différentes priorités scientifiques de l'unité, ainsi qu'une publication et une appli web venant illustrer nos recherches sur le Covid-19. Celui de l'équipe l'équipe OPE se compose de 4 productions (fiches de reconnaissance, jeu de données, bulletin mensuel et poster). Tous ces éléments sont disponibles sous forme de liens hypertextes ; les publications sont également disponibles sous forme d'un fichier .zip fourni en annexe.

Equipe Recherche

#Geolearning. La méthode d'inférence bayésienne INLA (Rue et al 2009) a simplifié l'implémentation et l'estimation de modèles de régression pour la modélisation prédictive et explicative de données complexes, volumineuses et souvent non gaussiennes, en intégrant des champs gaussiens dans les prédicteurs, par exemple pour tenir compte de dépendances spatiales. Dans [\[PUB-OHB18\]](#), cette méthode a été adaptée à la modélisation des dépassements de seuil, l'une des principales approches statistiques issues de la théorie des valeurs extrêmes.

#Epi-stat. L'article [\[PUB-CMS21a\]](#) concerne l'estimation des liens directionnels et pondérés d'un réseau géographique à partir d'un échantillon de trajectoires. Cet article inclut une description mathématique de la méthodologie et une application à la reconstruction des réseaux troposphériques. Il témoigne d'un corpus de publications [PUB-LMC18, Choufany (2020, thèse), PUB-CMS21b, Radici et al 2022] et d'une application web prochainement disponible permettant aux utilisateurs de générer leurs propres réseaux troposphériques.

#Evol-sys. L'article [\[PUB-RPB20\]](#) décrit analytiquement la trajectoire d'adaptation d'une population dans des environnements changeants en temps, en proposant une méthode de résolution d'EDP structurées en traits. Cette méthode permet de retrouver et de généraliser un ensemble de résultats obtenus par des mathématiciens de premier plan (e.g. Lorenzi et al 2015, Alfaro et al 2017, Figueroa Iglesias & Mirrahimi 2019, Carrère & Nadin 2020, Calvez et al 2021).

#Obs-Model. Les données opportunistes sont entachées de nombreux biais dont il est difficile de se défaire du fait de la méconnaissance de l'effort d'échantillonnage. Dans le travail [\[PUB-BJB21\]](#), nous proposons une méthode pour estimer conjointement la distribution des espèces et l'effort d'échantillonnage afin de corriger en partie ces biais. Nous appliquons notre approche sur des données issues des sciences citoyennes récoltées via l'application Pl@ntNet.

#Spread. L'approche 1D-2D [\[PUB-ZSO21\]](#) a permis de décrire la dynamique d'un ravageur et de son prédateur naturel sur un paysage agricole réel et d'étudier le rôle de la structure du paysage sur l'efficacité du biocontrôle. Ce travail fait appel aux compétences de l'équipe en simulation numérique, en modélisation et en écologie.

Covid-19. Les travaux menés sont révélateurs du savoir-faire de l'équipe en matière de couplage modèles mécanistes-données et de mise en place d'applis web pérennes. L'article [\[PUB-RKP20a\]](#) a été cité une centaine de fois (google Scholar). L'ensemble d'[applications R-Shiny](#) fonctionne sans interruption depuis début 2020 à partir d'extractions automatiques des données JHU-CSSE.

Equipe OPE

Fiche de reconnaissance de *Spodoptera frugiperda* [\[COL-OPE21a\]](#). Ce type de fiche est produit par la Plateforme ESV pour chaque organisme nuisible réglementé ou émergent et mis à disposition du public et des professionnels de terrain qui réalisent les observations, piégeages et prélèvements sur l'ensemble du territoire.

[Dataset de *Xylella fastidiosa*](#). L'une des missions de la Plateforme ESV est de mettre à disposition, à des fins de recherche, des jeux de données de surveillance officielle consolidés et anonymisés, et ainsi favoriser des interactions entre missions opérationnelles de surveillance et recherche.

[Bulletin mensuel n°37](#) (Décembre 2021). Les bulletins mensuels sont des productions issues de l'activité de veille internationale de la plateforme. Ce bulletin mentionne notamment des signalements de nouveaux cas en Europe de pathogènes affectant la tomate et la vigne.

[Poster XF-Actors Meeting \(2019\)](#) "Risk maps and risk-based surveillance for early detection of *Xylella fastidiosa* in continental France". Ce poster présenté à l'occasion d'une conférence européenne sur *Xylella fastidiosa* décrit une approche pour cartographier le risque de présence de *Xylella* en France et concevoir une stratégie de surveillance fondée sur le risque.

Références bibliographiques

Alfaro, M.; Berestycki, H. & Raoul, G. *The effect of climate shift on a species submitted to dispersion, evolution, growth, and nonlocal competition*. SIAM Journal on Mathematical Analysis, 2017, 49, 562-596.

Calvez, V.; Henry, B.; Méléard, S.; & Tran, V. C. *Dynamics of lineages in adaptation to a gradual environmental change*. arXiv preprint arXiv:2104.10427, 2021.

Carrère, C. & Nadin, G. *Influence of mutations in phenotypically-structured populations in time periodic environment*. Discrete & Continuous Dynamical Systems-B, American Institute of Mathematical Sciences, 2020, 25, 3609.

Figueroa Iglesias, S. & Mirrahimi, S. *Selection and mutation in a shifting and fluctuating environment*. HAL Preprint 02320525, 2019.

Foutel-Rodier, F. & Etheridge, A. M. *The spatial Muller's ratchet: surfing of deleterious mutations during range expansion*. Theoretical Population Biology, 2020, 135, 19-31.

Garnier, J.; Giletti, T. & Nadin, G. *Maximal and minimal spreading speeds for reaction diffusion equations in nonperiodic slowly varying media*. Ann Inst Henri Poincaré, 2012, 11, 173-188.

Lorenzi, T.; Chisholm, R. H.; Desvillettes, L. & Hughes, B. D. *Dissecting the dynamics of epigenetic changes in phenotype-structured populations exposed to fluctuating environments*. Journal of Theoretical Biology, 2015, 386, 166-176.

Radici, A.; Martinetti, D. & Bevacqua, D. *Early-detection surveillance for stem rust of wheat: insights from a global epidemic network based on airborne connectivity and host phenology*. Environmental Research Letters, 2022, in press.

Roques, L.; Garnier, J.; Hamel, F. & Klein, E. K. *Allee effect promotes diversity in traveling waves of colonization*. Proc Natl Acad Sci USA, 2012, 109, 8828-8833.

Rue, H., Martino, S., & Chopin, N. *Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations*. Journal of the royal statistical society: Series B, 2009, 71, 319-392.

Soubeyrand, S., & Roques, L. *Parameter estimation for reaction-diffusion models of biological invasions*. Population Ecology, 2014, 56(2), 427-434.