

MIELLEES - Le Système Informatique (SI) MIELLEES, un outil de mutualisation des données des balances connectées de ruches pour décrire les miellées

Allier F.^{1,2}, Aupinel P.^{2,3}, Brun F.⁴, Chevallereau C.^{2,3}, Dangléant A.^{1,2}, Frontero L.^{2,5}, Jourdan P.^{2,6}, Kretzschmar A.⁷, Labouyrie M.^{1,2}, Laperruque F.^{2,8}, Maisonnasse A.^{2,6}, Mugnier H.⁹, Odoux J-F.^{3,10}, Serard P.^{1,2}

¹ ITSAP-Institut de l'abeille, 228 route de l'aérodrome, F-84914 Avignon cedex 9

² UMT Prade, 228 route de l'aérodrome, F-84914 Avignon cedex 9

³ INRAE, UE1255, UE APIS, F-17700 Surgères

⁴ ACTA, 149 rue de Bercy, F-75595 Paris cedex 12

⁵ ADANA, Maison de l'agriculture, 55 av. Cronstadt, BP 279, F-40005 Mont-de-Marsan cedex

⁶ ADAPI, Maison des Agriculteurs, 22 Avenue H. Pontier, F-13626 Aix en Provence

⁷ INRAE, Unité Biostatistique et Processus Spatiaux, Site Agroparc, F-84914 Avignon

⁸ INRAE – UMR 1388 Genphyse, F-31326 Castanet Tolosan

⁹ ADTP, 1 Avenue du capitaine Anjot, F-74960 Annecy

¹⁰ INRAE-UNICAEN EVA-UMR 950, Écologie des prairies, Université de Caen, F-14032 Caen cedex

Correspondance : fabrice.allier@itsap.asso.fr

Résumé

Les balances automatiques et connectées mesurant le poids des ruches au fil des miellées sont utilisées par les apiculteurs, comme par les acteurs de la recherche et du développement, afin d'acquérir des informations en temps réel sur la progression de la miellée. Ces informations sont diffusées chez les apiculteurs ou les fournisseurs de services et constructeurs de balances mais non capitalisées. Le travail réalisé pendant le projet Casdar MIELLEES (2017-2020) montre qu'il est possible de créer une chaîne d'acteurs compétents et pluridisciplinaires pour développer des outils afin de mieux utiliser et valoriser des données numériques issues des balances connectées des apiculteurs. Un système informatique (SI) appelé le SI MIELLEES a été réalisé. Il collecte, structure, prépare et développe spécifiquement des séries temporelles pour les analyser grâce à des routines informatisées. Des prototypes d'applications de services ont été conçus et testés pour la valorisation des références acquises.

Mots-clés : production de miel, colonies d'abeilles domestiques, science participative, balances connectées, données numériques

Abstract: SI MIELLEES - a tool to share data from connected hives scales to describe honeydews

The automatic and connected scales measuring the weight of the hives during the honeydew are used by beekeepers, and by research and development actors, in order to collect real-time information on the state of the honeydew. This information is shared among beekeepers or service providers and scale manufacturers, and not capitalized. The work carried out during the Casdar MIELLEES project (2017-2020) showed that it is possible to bring together a chain of competent and multidisciplinary actors to develop tools and better use and enhance digital data from the connected scales of beekeepers. A computer system called SI MIELLEES was created. It collects, structures, and prepares time series for analysis using computerized routines developed specifically. Prototypes of service applications were designed and tested to enhance the value of the collected data.

Keywords: honeydew, honeybee colony, participatory science, connected scales, digital data.

Introduction

De plus en plus d'apiculteurs sont à la recherche de références ou d'outils techniques pour les aider dans leurs prises de décisions quotidiennes. Cela permet de positionner leurs stratégies et leurs performances technico-économiques parmi des référentiels établis et un ensemble d'informations plus larges. Actuellement, les données générées sur une exploitation par un apiculteur sont peu partagées et ne sont donc que peu valorisées au-delà du cercle proche de l'apiculteur lui-même ou éventuellement dans un cadre associatif et structuré. Ces données et les multiples expériences acquises au quotidien sur les exploitations, lorsqu'elles sont rendues accessibles, favorisent les échanges techniques et nourrissent l'innovation. Hambleton avec son équipe s'en était aperçu en 1925 aux Etats-Unis, lorsqu'il réalisa ses expériences et affirmait à son époque, et sans doute de même aujourd'hui, qu'il n'y avait pas de méthode pour prédire si une localité sera profitable à l'apiculteur si ce n'est par les résultats obtenus par d'autres apiculteurs.

L'usage des capteurs, à fonctionnement électromécanique puis numérique a attiré la curiosité des chercheurs dès la première moitié du siècle dernier (Decourtye et al., 2018). En 1925, Hambleton rapportait par exemple que les enregistrements de pesées horaires de ruches de colonies d'abeilles domestiques étaient rares. Un autre auteur (Dufour, 1897) a souligné la valeur des pesées horaires et a montré à quel point les variations horaires des poids tout au long de la journée pouvaient varier énormément les jours présentant le même gain net. Dans les mêmes années, Lundie (1925) reportait pour la première fois la construction et l'utilisation d'un compteur électromécanique d'abeilles sortant de la ruche. En France l'engouement pour les objets connectés, resté jusqu'alors relativement confidentiel dans les cercles scientifiques, débute au début des années 2000 avec la balance électronique. En effet, il est reconnu que le poids de la ruche est un indicateur de la quantité de miel stockée (Decourtye et al., 2018). Nous assistons maintenant à une effervescence d'idées, de solutions et de capteurs qui se proposent de venir appuyer voire faciliter les pratiques des apiculteurs, mais également de mieux révéler le fonctionnement des colonies, de leur capacité à se développer dans un milieu donné et du rendement en miel attendu, ou être des révélateurs « bio-indicateurs » de l'état de l'environnement (Decourtye et al. 2018). La balance connectée devient le premier outil pratique qui offre à l'apiculteur une nouvelle « visibilité » à distance sur son rucher ou d'une colonie. Les frais de déplacement pour réaliser la transhumance des ruches pour la production de miel représentent le principal poste de dépenses, environ 40 % des coûts de production (Aulanier et Ferrus, 2018). L'investissement dans une balance interrogeable à distance intervient logiquement dans une stratégie de diminution des charges de déplacements tout en maintenant le flux d'information par le capteur sur l'évolution du gain de poids d'une ruche positionnée à plusieurs dizaines de kilomètres du siège d'exploitation.

L'idée de mieux valoriser les données issues de ces balances pour les apiculteurs et la recherche a fondé la création d'un partenariat autour du projet MIELLEES : Collecter, organiser, analyser et restituer les données de poids issues des balances connectées des ruches. L'article dresse, dans une première partie, le panorama des acteurs associés à la valorisation de ces données ainsi que leurs attentes. Il décrit techniquement dans un second temps l'outil créé appelé Système Informatique MIELLEES. Enfin, nous présentons les étapes testées successivement pour gérer et préparer les données, et quelques applications prototypes pour valoriser ces données numériques.

1. La mutualisation des données individuelles : une meilleure valorisation de l'information et compréhension des problématiques collectives

1.1 *La transition vers un usage collectif des données numériques des balances*

1.1.1 Avènement des balances connectées en apiculture

Dès la fin d'une floraison emblématique ou la saison apicole passée, chaque producteur de miel tente de se faire une idée de sa production au regard de celles de ses pairs lors des rencontres spontanées ou

organisées, par téléphone ou email, et plus récemment au travers des réseaux sociaux illustrés de photos ou de vidéos. Depuis toujours, les agriculteurs ont été à la recherche de références techniques et économiques pour se comparer et s'améliorer. Les producteurs de miel n'échappent pas à cette règle. Ces informations les aident dans leurs prises de décisions quotidiennes. De ces échanges, il reste peu de prise de note pour déduire les tendances et tirer des bilans. Pour obtenir une idée plus précise de la production de miel en France une année donnée, les pouvoirs publics et les structures de la recherche et développement (R&D) de la filière apicole ont pris l'habitude de créer des enquêtes, relativement chronophages. Elles sont aujourd'hui largement diffusées sur internet et les résultats sont disponibles plusieurs mois après, pendant la saison suivante.

L'avènement et l'accessibilité des objets connectés en agriculture ouvrent la voie vers une transition de la gestion, du stockage et de la valorisation des données numériques. Un accès plus rapide, voire direct, à ses données est désormais possible ainsi qu'aux indicateurs et références qu'elles permettent d'élaborer. Si des habitudes sont déjà prises dans d'autres filières, l'apiculture n'y échappe pas, bien qu'elle en soit à ses débuts. Parmi ces innovations, la balance électronique, arrivée au début des années 2000, devient le premier outil pratique qui offre à l'apiculteur une nouvelle « visibilité » du comportement des colonies de son rucher, et cela à distance. Dans ce marché de la balance connectée, encore jeune et instable, sur la trentaine d'entreprises identifiées au début du projet, une douzaine semblait fournir régulièrement les apiculteurs français¹ et cet équipement bénéficie, depuis plus de 15 ans, des subventions du Programme apicole européen, gérées par FranceAgriMer. Le nombre de balances commercialisées en France peut être estimé à quelques milliers d'exemplaires.

Les initiatives collectives pilotées par des Associations de Développement de l'Apiculture (ADA) visent à partager plus largement les données générées par des balances connectées (Figure 1). Les balances peuvent être prêtées à des apiculteurs ou être intégrées à un protocole expérimental de suivi de miellées². Techniquement, les apiculteurs consultant ces données accèdent aux périodes de floraison des espèces ciblées, le déclenchement des miellées, les dynamiques de gain de poids de plusieurs colonies d'un même rucher ou appréhender la fin d'une miellée. Ces données leur permettent en outre de positionner leurs stratégies et leurs performances technico-économiques parmi des référentiels établis.



Figure 1 : Le dispositif ECOBEE (INRAE APIS) avec installation de balances du constructeur MicroEL

1.1.2 L'usage des balances par les apiculteurs

En 20 ans, la balance connectée s'est imposée comme un outil technique largement utilisé dans les exploitations. Nombreux sont les apiculteurs qui jugent de la pertinence de se déplacer pour intervenir, principalement à partir de l'information obtenue d'une ou plusieurs ruches posée(s) sur balance(s) connectée(s). Une enquête réalisée en 2017 dans le projet MIELLEES auprès d'un réseau de 30 apiculteurs professionnels renforce cet intérêt premier exprimé de gain de temps. Ainsi, l'apiculteur

¹ <https://itsap.asso.fr/outils/balances-automatiques/>

² Voir les sites internet de ADAPI, ADANA, ADA Grand Est, ADA Occitanie, ADA Bretagne, ADA AURA

décidera de visiter ses colonies sur la base d'un premier diagnostic de la situation à distance. Les préventions d'essaimage, les ajouts ou retraits de hausses, le nourrissage (apport complémentaire de sucre), les périodes de disette alimentaire sont alors mieux anticipés.

Pour cet échantillon d'apiculteurs, deux marques étaient les plus utilisées : CAPAZ (constructeur allemand historique des années 2000), et Optibee (CAD Création) commercialisée au début des années 2010. La marque Beewise (modèle français historique) est utilisée par 5 apiculteurs de cet échantillon. Le quatrième constructeur est Beeguard dont deux apiculteurs de cette enquête étaient équipés.

Globalement, ces outils sont particulièrement appréciés pour la fiabilité de connexion et l'envoi régulier de données, même si quelques dysfonctionnements apparaissent (les connexions réseaux le plus souvent) selon les dires d'apiculteurs et les modèles. L'interface de lecture des données est citée comme un point positif sans être majoritaire dans les citations et devient un vrai support de services pour les constructeurs. Les apiculteurs portent cependant un jugement plus nuancé sur le poids des modèles les plus lourds, jugés aussi trop encombrants. L'autonomie et l'entretien peuvent apparaître à la marge comme un point négatif (cités par 5 apiculteurs sur 30 interrogés). Ce sujet est un point d'amélioration constant pour les constructeurs. Selon les balances et leurs options, les données disponibles sont différentes, mais la plupart des modèles observés fournissent le poids et des données météorologiques, souvent grâce à des capteurs externes à la ruche (humidité, pluviométrie, température). Le poids est la variable la plus concrète et pour laquelle les apiculteurs portent le plus d'intérêt.

L'enquête s'intéressait aussi à connaître le positionnement des apiculteurs dans la mise à disposition de leurs données. Si pour une majorité, le partage de leurs données à des fins de recherche et développement dans des conditions d'anonymat ne semble pas être une contrainte forte, trois points peuvent cependant limiter leurs engagements : la géolocalisation précise des balances, les niveaux de partage et de confidentialité de la donnée.

1.2 Charte d'engagement des acteurs dans le projet MIELLEES

La création d'un système informatique pour organiser la collecte, le stockage et l'analyse des données est une innovation pour la filière apicole et les acteurs privés fournisseurs de balances connectées. Pour être opérationnelle, cette chaîne de nouveaux partenaires doit détenir des compétences diverses pour rassembler et analyser des données mais également pour gérer les droits de propriété et de communication de ces données numériques (ACTA, 2016). L'implication de nouveaux acteurs de la recherche et du développement libère la concentration des informations chez les propriétaires ou chez quelques opérateurs. Elle offre une nouvelle envergure et des perspectives valorisantes de ces données pour l'ensemble des acteurs de la filière apicole.

Pour les apiculteurs, cette chaîne de partenaires représente une opportunité de s'impliquer dans une démarche de sciences participatives en donnant accès à leurs données. Ils sont en attente de pouvoir visualiser les performances de leurs colonies et de les comparer à celles des colonies voisines ou de leurs collègues, situées dans un même contexte de miellée. A plus long terme, et sous réserve d'accéder à des jeux de données complets, c'est-à-dire dont la miellée visée et la géolocalisation sont connues, des questions plus complexes de recherche peuvent être adressées et des recommandations techniques apportées en retour aux apiculteurs sous forme d'indices, d'indicateurs et d'outils d'aide à la décision.

Pour les constructeurs de balances, leurs compétences et savoir-faire ainsi que leur proximité aux apiculteurs en tant que fournisseurs de matériel, les positionnent au cœur du réseau partenarial créé pour construire collectivement un système informatique. De la fabrication de l'outil connecté à la gestion du flux des données et jusqu'à la mise en forme de cette donnée pour le propriétaire de la balance, leurs attentes sont fortes pour participer à la valorisation des connaissances acquises auprès de leurs clients apiculteurs, associations ou entreprises. Pour les constructeurs, leurs services peuvent être améliorés et

complétés de nouvelles informations et indicateurs analysées par les acteurs de la recherche et restituées aux apiculteurs via leurs applications de service.

Pour les structures d'accompagnement technique et de développement telles que les ADA, les intérêts sont multiples. A ce jour, les données générées par les balances sont visualisées en temps réel principalement, peu exploitées et stockées pour des analyses ultérieures. Leurs besoins se traduisent à des niveaux collectifs et individuels. Les conseillers apicoles sont par exemple amenés à expertiser des stratégies pour des exploitations individuelles à partir de la connaissance des dynamiques des colonies, des évolutions de poids au cours d'une saison, du niveau de la production et des rendements. Ils créent également des bilans de production de miel de fin de saison ou encore ils étayent techniquement et économiquement les demandes de reconnaissances en calamités agricoles.

Pour les structures de Recherche, Développement et d'Innovation telles que INRAE³, l'ITSAP⁴, l'intérêt est d'accéder à des données individuelles brutes pour les valoriser autour de questions de recherches fondamentale et appliquée utiles à la compréhension de phénomènes complexes et à la production d'outils d'aide à la décision (OAD). L'analyse de ces grandes quantités de données contextualisées peut améliorer la connaissance des productions de miel par type de territoire et par période de miellée pour participer à la création d'observatoires régionaux et national de la production de miel. D'un point de vue plus fondamental, les séries temporelles valides peuvent intégrer des analyses pour modéliser des situations observées par les apiculteurs et en comprendre les causes et les effets. Les fortes mortalités et affaiblissements des colonies pourraient être étudiés avec ces nouveaux jeux de données.

Même si les attentes de chacun semblent converger et que chaque acteur puisse et doive trouver sa place, rassembler tous ces acteurs au sein d'un nouvel écosystème d'acteurs nécessite de s'interroger collectivement sur de nouvelles questions auxquelles sont confrontés tous les acteurs :

- Que signifie valoriser les données et comment sont-elles sécurisées ?
- Qui en est propriétaire et comment les droits d'accès aux données sont-ils gérés ?
- Faut-il favoriser la circulation des données vers un système collectif de stockage et d'analyse ?
- Quelles sont les possibilités de création de nouvelles connaissances grâce à ces données ?

Pour y répondre, et plus globalement à l'échelle nationale, les acteurs s'organisent et proposent leurs points de vue au travers de chartes (FNSEA - Jeunes Agriculteurs, 2018) ou livres blancs (ACTA, 2016 ; Renaissance numérique, 2018). Dans le cadre du projet MIELLEES, et pour mieux trouver les réponses à ces questions, une charte d'engagement a été rédigée afin de présenter notamment le contexte, les objectifs et les intérêts d'être partie prenante d'une telle démarche pour des données apicoles générées à partir des balances connectées. Cette dernière a aussi vocation de rassurer et répondre aux questions sur la sécurisation, l'usage et l'accès aux données. Elle présente également les services apportés par le consortium MIELLEES et les engagements attendus des parties afin que le système soit équilibré et ne déstabilise pas un acteur. L'objectif de ce partenariat est finalement de sortir des sentiers battus, en passant d'un usage quasi instantané ou éphémère et individuel de la donnée, à un usage confiant, partagé et durable des données numériques. Les partenaires se sont entendus sur les points suivants dans le cadre de la charte :

- Intérêt collectif et partagé autour des enjeux portés par le projet MIELLEES : utilité de rassembler les données brutes (non lissées sur une période) issues des balances, de les sécuriser, les organiser et les analyser. La tâche consiste ainsi à organiser la mise à disposition auprès des acteurs techniques et de la recherche pour mener des travaux d'analyses collectifs, produire des référentiels et des outils prédictifs, puis les restituer aux apiculteurs.

³ Institut national de recherche pour l'agriculture, l'alimentation et l'environnement

⁴ Institut technique et scientifique de l'apiculture et de la pollinisation

- Pertinence partagée de créer un lieu d'échange sur des sujets communs autour des balances et des données générées. Positionner dans cette nouvelle organisation, l'apiculteur comme fournisseur consentant au départ et propriétaire de la donnée et "client" ou bénéficiaire final de l'information traitée.
- Consensus de ne pas déséquilibrer les offres existantes de services produites par les constructeurs (interfaces de visualisation internet). L'analyse des données rendues accessibles devra apporter des plus-values grâce à la diversité des compétences issues des structures commerciales privées et des partenaires du R&D.
- Besoin de renforcer le SI MIELLEES avec des nouvelles fonctions facilitant l'accès aux données par des requêtes pour les diverses structures de la chaîne de valeur : Instituts de recherche pour la réalisation des analyses approfondies des données ; constructeurs pour la valorisation sur leur propre interface web des données analysées ; structures de développement apicole pour la production de suivis de campagne de leurs adhérents et de conseil technique.

2. Le Système Informatique MIELLEES

L'objectif du projet est de permettre à tous les apiculteurs de pouvoir bénéficier à terme du service qu'apportera le SI MIELLEES, en proposant ce partenariat à plusieurs constructeurs. Dans le temps imparti du projet, une longue période a été consacrée au recensement des constructeurs existants et commercialisant des balances connectées en France, ainsi qu'aux discussions sur leur place dans ce nouveau partenariat.

2.1 Identification des constructeurs de balances

Une trentaine de structures ou de projets, constructeurs, concepteurs de balances, ont été identifiés en 2017 dès le début du projet. Compte-tenu de ce nombre élevé, un choix arbitraire n'avait pas été décidé, mais plutôt de les informer de l'initiative, de les associer à la réflexion et finalement, de démarrer avec les structures les plus intéressées.

A partir de ces structures, une première liste de 16 constructeurs ayant une commercialisation active des balances au cours de l'année 2017, a été retenue. Des demandes de compléments techniques simples ont permis d'aboutir à un descriptif de 12 modèles de balances connectées, accessible sur une page du site internet de l'ITSAP⁵. Sur cette douzaine, seulement six d'entre eux ont montré un intérêt à la dimension R&D présentée par les partenaires. Parmi ces 6 constructeurs, trois sont bien présents sur le marché français à savoir CAD Création (Optibee), Connected Beekeeping (Beescale) et Beeguard. Le constructeur croate MicroEL, non présent en France, fut fournisseur de balances d'un partenaire, l'Unité expérimentale INRAE APIS (17). Le constructeur LabelAbeille intéressé ne commercialise que très peu d'outils connectés aux apiculteurs professionnels. Seules les balances des marques Optibee, Beeguard et CAPAZ étaient utilisées par le groupe d'apiculteurs volontaires adhérents des deux structures régionales de développement de l'apiculture impliquées, l'ADANA⁶ et l'ADAPI⁷. Le constructeur historique allemand CAPAZ, très présent en France, n'a pas souhaité s'investir davantage et a quitté les discussions avant les deux ateliers de concertation et de co-construction organisés en 2018 avec les 5 constructeurs finalement engagés. En fin de projet, seuls les 4 constructeurs Beeguard, Optibee, Connected Beekeeping et LabelAbeille avaient pu adapter leur serveur pour l'envoi des données selon le format unique retenu, et des tests réalisés de bon fonctionnement ont été effectués. Le constructeur croate MicroEL n'a finalement pas été en mesure d'adapter ses serveurs à ce cahier des charges commun.

⁵ <https://itsap.asso.fr/outils/balances-automatiques/>

⁶ Association régionale de développement de l'apiculture en Nouvelle aquitaine

⁷ Association régionale de développement de l'apiculture en Provence

2.2 Création d'une passerelle numérique

Le recueil des données des balances dans le SI MIELLEES a entraîné le développement d'une passerelle numérique pour aiguiller ces données depuis les serveurs collecteurs des constructeurs des balances vers une unique base de données. Bien que les principales variables collectées soient les mêmes quels que soient les modèles de balances (poids, dates et heures), elles sont émises dans des formats et à des fréquences hétérogènes. Elles nécessitent donc de passer par une uniformisation de leur format. Afin de réaliser cette tâche plusieurs travaux intermédiaires ont été réalisés.

Courant 2018, un recueil des spécifications techniques des balances commercialisées a été demandé aux 6 constructeurs intéressés et a abordé les quatre thèmes suivants :

- Caractéristiques générales (dynamique, précision, types de capteur supplémentaire, fréquence de mesure...)
- Flux de données (type d'envoi, format, technologie...)
- Métrologie et garantie (étalonnage, certification, garantie du produit...)
- Marchés (nombre de balances, maturité, vision marché, vision projet...)

Ce travail a permis d'appréhender la diversité des balances et la manière dont les données sont fournies ; d'établir une base technique et qualitative pour engager des échanges plus approfondis avec les constructeurs ; et de préparer le cahier des charges pour la mise en place de la passerelle.

2.3 Cahier des charges du SI MIELLEES

En 2020, en fin de projet, l'ensemble des partenaires incluant les constructeurs se sont entendus sur le cahier des charges du SI MIELLEES. Techniquement, deux aspects incontournables ont été ajoutés, à savoir la capacité à intégrer des grandes quantités de données par lot, ainsi qu'une absolue confidentialité des informations remontées dans la base. Par exemple, aucun constructeur ne doit pouvoir accéder aux données d'un autre constructeur, pour des raisons de concurrence évidentes et les apports de données, via une procédure automatisée, doivent être authentifiées et chiffrées. Le cahier des charges précise en outre le contexte des éléments qui seront enregistrés dans la base de données selon un format unique quel que soit le constructeur engagé (Tableau 1). Ces données sont stockées sur des serveurs hébergés à l'INRAE en France, non accessibles sans autorisation et leur communication est nécessairement anonyme et la géolocalisation floutée, conformément à la réglementation générale sur la protection des données (RGPD).

Tableau 1 : Description des mesures et formats validés pour chacun des éléments intégrant le SI MIELLEES. Les noms et coordonnées des propriétaires des balances ne sont pas communiqués.

Mesure	Code normalisé	Unité	Type	Présentation de la donnée (exemple)
Identifiant du constructeur	const		3 caractères	OPT
Identifiant unique de la balance	bal		Chaîne de car.	Bal_ruche1
Identifiant de la ruche	ruche		Chaîne de car.	r154
Pesée de la balance	poids	gramme	Entier	62250
Qualification de la géolocalisation	qloc		Chiffre	1
Latitude	lat	degrés décimaux	Chaîne de car.	"43.6991523910"
Longitude	lon	degrés décimaux	Chaîne de car.	"-5.24579985"
Activité	activ		Chaîne de car.	"lavande"
Timestamp en ns depuis le 01/01/1970 UTC		ns	Entier	1,58636E+17

A ce jour (fin 2021), les données d'environ 150 balances circulent vers le SI MIELLEES, en grande majorité issues des balances gérées par les ADA. Les sorties opérationnelles (indicateurs, références) à

valoriser auprès des apiculteurs sur les applications des constructeurs, n'ont pas dépassé le stade de maturité de prototypage et ne sont donc pas fonctionnelles en routine. Elles seront à construire ultérieurement, après l'automatisation complète des routines et procédures d'analyses des données décrites dans la troisième partie de cet article.

2.4 Application de requêtes et d'analyses pour les structures de R&D

Pour assurer les travaux des chercheurs, un prototype d'application internet de requête, de gestion, d'analyses et de tests des références a été conçu et testé. Cette application (Figure 2) importe les séries temporelles disponibles, nettoyées et corrigées depuis le SI MIELLEES. Elle est programmée selon le langage MYSQL, compatible avec d'autres outils développés par l'ITSAP et l'INRAE pour les expérimentateurs et conseillers apicoles du réseau (Dangléant, 2021). Celle-ci est fondée sur des référentiels « métiers » définis pour structurer la saisie et l'importation de données (mesures, unités, structure, profil des structures de recherche, activité des balances (miellées ciblées, sous activité). Cette interface permet aux conseillers apicoles de suivre les évolutions des dynamiques des colonies sur les périodes étudiées à partir de leurs propres choix de requêtes.

Graphique multi mesures

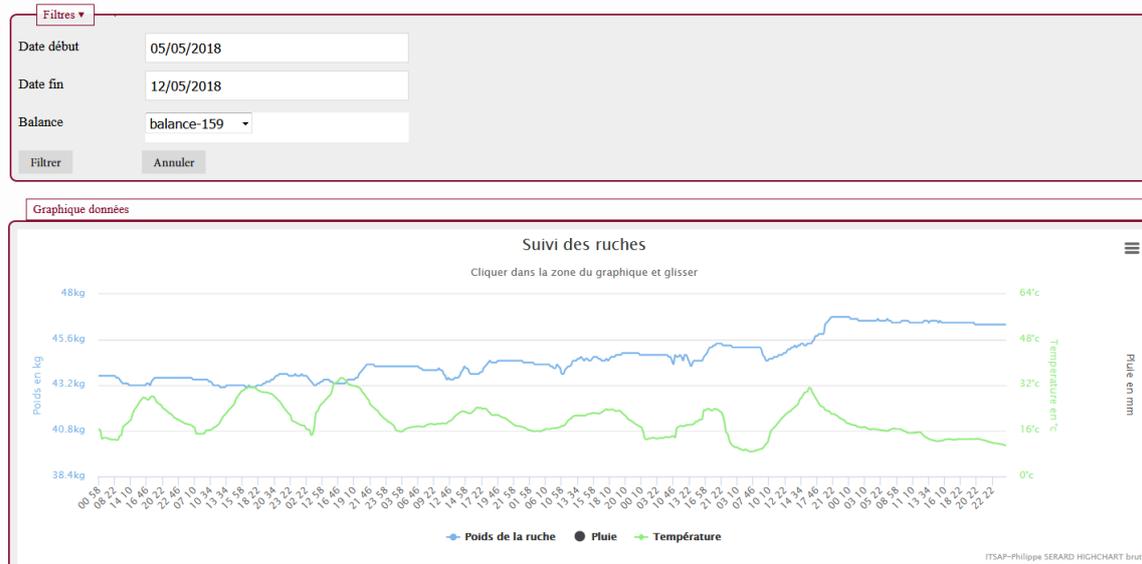


Figure 2 : Illustration de l'interface de saisies, d'importation et d'affichages de données à partir du SI MIELLEES pour les conseillers apicoles.

3. Gestion des séries temporelles collectées et création d'applications de valorisation des références

3.1 Recherche de séries temporelles de données valides

Avant les étapes d'analyses statistiques, une prise en main et un nettoyage des données constituent une étape indispensable pour créer des ensembles de séries temporelles d'une saison apicole par exemple pouvant être considérées comme propres ou analysables. Plusieurs tâches sont nécessaires et décrites succinctement ci-après.

3.1.1 Repérage des accidents de pesées et interventions apicoles sur les ruches

La prise de mesures par une balance peut générer des accidents de pesées, résultant en un pic de poids de plusieurs dizaines de kilogrammes en l'espace de quelques minutes ou d'une intervention apicole (pose / retrait de hausse). Ces accidents et ce type d'interventions apicoles, visibles sur la Figure 3, sont

par conséquent à l'origine de variations de poids importantes enregistrées par la balance sur des laps de temps courts (deux à trois mesures), et seront supprimés de la série de mesures.

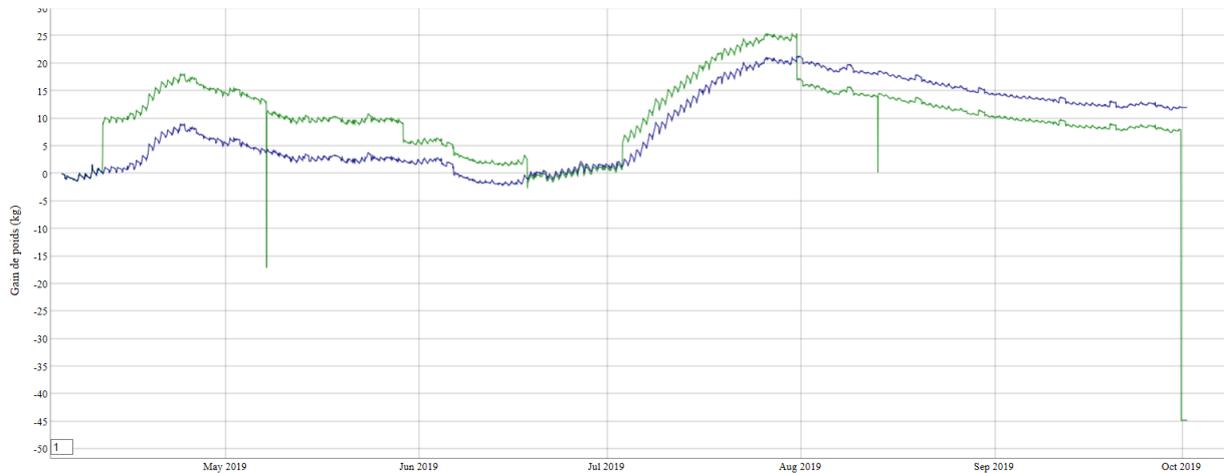


Figure 3 : Représentations graphiques des courbes de poids avec la méthode de correction des données. En vert, la courbe représente l'évolution du poids incluant les interventions de l'apiculteur (ajout/retrait de hausses, visite de colonie) et les artefacts de l'électronique. En bleu, la courbe représente l'évolution du poids corrigé des interventions telles que les ajouts/retrait de hausses, visites de colonies.

3.1.2 Correction des données de pesée, données manquantes et segmentation des séries

Afin d'établir une lecture de la dynamique de poids associée à la vie de la colonie seule, c'est-à-dire débarrassée des accidents de pesée et des interventions apicoles, une routine de correction des données de pesées mesurées a été créée. Ces perturbations sont d'abord repérées par une procédure simple. Un intervalle de poids considéré comme biologiquement cohérent est établi tel que d'une mesure de poids à la suivante, sur une heure, un gain ou une perte de poids de plus ou moins 2 kg représentent une variation de poids tolérée. En dehors de cet intervalle de poids, les valeurs sont corrigées en repositionnant les points de mesures venant après la variation, au niveau du point de mesure précédent cette variation. Pour chaque ligne du tableau de données de pesée, si la différence de poids entre deux mesures consécutives est hors de l'intervalle de poids toléré, alors la mesure à corriger est remplacée par le dernier poids corrigé accepté (Labouyrie, 2020).

De plus, dans le cas où les séries de données de poids comportent des données temporelles manquantes, une autre routine de contrôle a été mise en place. Celle-ci consiste à vérifier le pas de temps séparant deux mesures consécutives présentes dans le jeu de données, puis repérer les prises de mesures manquantes et à les ajouter artificiellement. L'étape suivante de segmentation des courbes sera ainsi moins faussée. La méthode de segmentation appliquée aux mesures de séries temporelles générées par les balances connectées permet de décrire les fortes variations d'une série de mesures. Lorsque certaines valeurs changent brusquement, elles sont appelées points de rupture (ou cassure). Afin de les détecter plus facilement, Muggeo (2017) a mis au point un package concernant la régression segmentée intitulé « segmented » et implémenté sous R (R Core Team 2013). Ce package d'analyse a été repris dans l'analyse du jeu de données grâce à la fonction *segmented.lm()* dans l'objectif de déterminer des périodes de miellées (Labouyrie, 2020).

3.2 Identification des périodes de miellées et établissement d'un calendrier apicole réaliste

L'étape suivante consistait à manipuler les séries temporelles segmentées afin d'identifier des périodes d'intérêt communes à l'ensemble des ruches sur lesquelles il est possible de quantifier un gain (les miellées) ou une perte de poids de la ruche (les récoltes ou des dynamiques biologiques de la colonie)

et d'y associer une activité de la colonie. Pendant le projet MIELLEES, ces travaux d'analyse se sont focalisés sur un jeu de données d'une trentaine de balances gérées par le partenaire INRAE APIS du Magneraud (17), pendant la période d'activité des colonies, de mars à fin septembre 2019. Des courbes représentant l'évolution des poids des ruches propres à la dynamique de la colonie ont été obtenues, en corrigeant les données de pesée. Puis la routine établie a permis de délimiter dans le temps les périodes de gain ou de perte de poids majeures grâce à la segmentation. Le résultat d'une telle procédure est semblable à celui présenté à la Figure 4. Ces périodes sont délimitées par deux dates, une date de début et de fin de période, généralement une miellée ou une saison, qui reflètent le fait qu'avant et après cette période, la dynamique de poids de la ruche est différente.

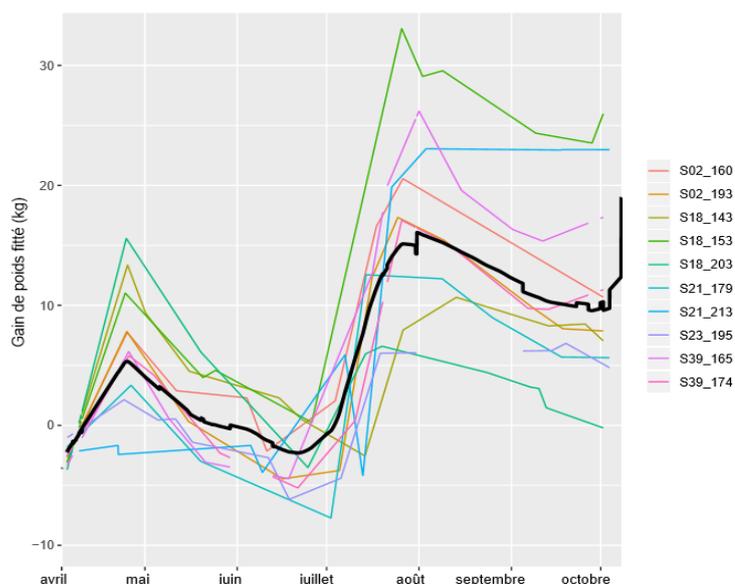


Figure 4 : Caractérisation des variations de poids, après segmentation, de 10 ruches suivies à l'INRAE APIS, sur une période d'avril à octobre. La courbe noire représente l'évolution moyenne du rucher.

3.3 Collecte des données de contexte

Le SI MIELLEES ne collecte aucune donnée de contexte, mis à part la géolocalisation de la ruche lorsque celle-ci est enregistrée par l'apiculteur ou qu'un capteur GPS est associé à la balance. La variable « Activité » précisant le nom de la miellée ciblée par l'apiculteur sur la période n'est pour l'instant pas ou mal renseignée par les propriétaires des balances. De ce fait, les analyses sont réalisées sans connaissance du propriétaire, du lieu de positionnement, ni des données de contexte. En dehors des cas expérimentaux, il n'est donc pas possible d'analyser les données au regard d'un contexte que ce soit un lieu ou une miellée commune sur une même période pour un ensemble de séries temporelles. Pour cela, les périodes de début et de fin de miellée potentielles ont été repérées à partir des courbes puis listées dans un tableau à compléter et envoyées au conseiller technique apicole, pour vérification et correction des dates, des miellées visées et de la géolocalisation de la balance, en charge du suivi des balances connectées dans les ADA partenaires du projet. Ces nouvelles données de contexte sont ensuite ajoutées au SI MIELLEES par une routine programmée afin de constituer un jeu de données complet et contextualisé.

Ces étapes ont abouti à la production d'un ensemble de routines qui constituent, sous forme de prototype actuellement, la procédure de prise en charge et de préparation des données brutes collectées dans le SI MIELLEES. L'analyse statistique de ces courbes et la création de nouvelles références constituera une étape à mener ultérieurement, tâche qui n'a pas été finalisée dans le cadre du projet Casdar MIELLEES

3.4 Identification et développement de services numériques d'aide à l'interprétation du suivi des miellées

Une autre tâche du projet consistait à créer des outils ou des applications prototypes pour valoriser et restituer les jeux de données préparés dans les tâches décrites précédemment. Nous dressons ci-dessous une liste des outils et procédures créés dont certains ont été testés.

- Nettoyage des données et intégration d'un champ d'annotation automatique des événements.
- Positionnement d'une performance d'une colonie par rapport à une référence (minimum, moyenne, maximum) d'une autre colonie ou d'un ensemble (Figure 4).
- Alerte de la date de démarrage d'une miellée sur une zone.
- Prédiction de l'évolution du poids dans les jours à venir.
- Création d'un indicateur météo en lien avec un comportement à prévoir de la colonie.
- Création d'un indicateur de l'attractivité et du risque d'un environnement agricole et forestier de la colonie.

Conclusion

Le Système Informatique MIELLEES a permis de constituer un partenariat technique nouveau en rassemblant des acteurs peu habitués à travailler ensemble et a abouti à la construction d'un outil prototype opérationnel. Les objectifs fixés par les partenaires étaient ambitieux et les nombreuses difficultés techniques, partenariales et financières ont limité la finalisation de certaines tâches, telles que l'analyse et la valorisation des données vers les apiculteurs. En effet, cet outil est composé de plusieurs entités, routines de programmation et applications qui nécessitent un travail de connexion entre elles afin de rendre le système fluide de l'entrée de la donnée brute à sa restitution aux apiculteurs sur les interfaces des constructeurs. Ce projet Casdar MIELLEES a ainsi posé les fondations techniques et partenariales d'une nouvelle approche de gestion de données numériques qui pourront être accueillies en grande quantité à l'avenir. Il ouvre un champ d'applications techniques dans le domaine de l'apiculture qui répondra à des attentes des apiculteurs comme l'optimisation de leurs interventions apicoles pour limiter leurs déplacements et améliorer le suivi des colonies. D'un point de vue scientifique, l'accès à des centaines de séries temporelles de poids des ruches sur plusieurs années offrira un terrain de jeu immense et nouveau pour aborder les questions de recherche différemment.

Depuis la fin 2020, les partenaires poursuivent le travail de maintenance et d'alimentation du SI MIELLEES pour permettre aux constructeurs d'accéder, à moyen terme, aux nouvelles références et les restituer aux apiculteurs.

Remerciements et financements

Le projet a été financé principalement par les fonds Casdar du ministère en charge de l'Agriculture et par les structures partenaires. Nous remercions les apiculteurs ayant participé aux différentes réunions, pour leur temps offert et leurs avis donnés également au travers des enquêtes. Ce projet a pu évoluer en tenant compte des entreprises de balances connectées et nous tenons à remercier leurs dirigeants pour le temps et l'expertise apportés. Enfin, les compétences de deux ingénieurs informaticiens, Aurélien Neveux et Jérôme Rousset, ont été indispensables dans le cadre du partenariat avec l'entreprise MIND-GIE Electronique et Développement.

Références bibliographiques

ACTA – Les instituts techniques agricoles, 2016. L'accès aux données pour la recherche et l'innovation en agriculture. Position des Instituts Techniques Agricole". PDF, 46 p.

Aulanier F., Ferrus C., 2018. L'apiculture professionnelle : aspects technico-économiques des exploitations. In: Decourtye A. (eds), Abeilles, des ouvrières agricoles à protéger. La France Agricole, ACTA Editions, Paris.

Dangléant A., Ricard F., Vidau C., La Perruque F., Pioz M., Guinemer M., Quintaine T., Decourtye A., 2021. IODA – Informatiser et Organiser les Données Abeilles. Innovations Agronomiques 82, 165-177 <https://doi.org/10.15454/1ygt-vw39>

Decourtye A., Dangléant A., Allier F., Alaux C., 2018. La ruche connectée : objet de surveillance environnementale, de zootechnie ou de découverte récréative. Innovations Agronomiques 67, 77-87. [dx.doi.org/10.15454/4LSDMM](https://doi.org/10.15454/4LSDMM)

Dufour L., 1897. Travail des butineuses et récolte du miel. In Apiculteur, vol. 41, No. 8, pp. 300-312, 6 figs. In Hambleton, 1925.

FNSEA - Jeunes Agriculteurs. 2018. Charte sur l'utilisation des données agricoles - DATA AGRI. Valoriser et sécuriser les données des exploitations agricoles dans les contrats

Hambleton J.I., 1925. The effect of weather upon the change in weight of a colony of bees during the honey flow. By, Apiculturist, in Charge of Bee Culture Investigations, Bureau of Entomology. DEPARTMENT BULLETIN No. 1339. United States Department of Agriculture.

Labouyrie M., 2020. Développement d'un outil statistique d'aide à l'analyse et à l'interprétation des miellées basé sur les données de balances automatiques (IOT). Rapport de stage, Master 2 Diplôme Ingénieur Agronome, Montpellier SupAgro

Lundie A.E., 1925. The flight activities of the honey bees, United States Department of Agriculture, Dept. Bull. No. 1328.

Muggeo Vito M.R. 2017. Interval estimation for the breakpoint in segmented regression: a smoothed score-based approach. Australian & New Zealand Journal of Statistics 59.3 (sept. 2017), p. 311-322. issn : 13691473. doi : 10. 1111/ anz.12200.

Renaissance numérique. 2018. La valeur des données en agriculture.

R Core Team 2013 R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria, 2013. url : <http://www.R-project.org/>.

Cet article est publié sous la licence Creative Commons (CC BY-NC-ND 3.0)



<https://creativecommons.org/licenses/by-nc-nd/3.0/fr/>

Pour la citation et la reproduction de cet article, mentionner obligatoirement le titre de l'article, le nom de tous les auteurs, la mention de sa publication dans la revue « Innovations Agronomiques », la date de sa publication, et son URL)