



SOLUTIONS
POUR
**NOTRE
AIR**



**Expérimentation
«Pollens» - Analyse
des données
d'inter-comparaison
de capteurs
automatiques de
comptage de pollens**

Juin 2023

RAPPORT TECHNIQUE FINAL

ANALYSE DES DONNEES D'INTER-COMPARAISON DE CAPTEURS AUTOMATIQUES DE COMPTAGE DE POLLENS



Référence FLUIDYN	0222029A
A destination de	 
Nombre de pages	44

Version	Date	Pages modifiées	Rédaction	Vérification
1.1	24/11/2022	--	Lobnat AIT-HAMOU	Amita TRIPATHI

Table des matières

LISTE DES ILLUSTRATIONS	4
LISTE DES TABLES	5
NOMENCLATURE	6
1 Contexte	7
2 Capteurs sur site.....	8
2.1 SYSTÈME SWISENSPOLENO MARS	8
2.1.1 Principe de mesure et sous-systèmes associés.....	8
2.1.2 Calibration et incertitudes intrinsèques (hors variabilité naturelle).....	10
2.2 SYSTÈME POLLESENSE	13
2.2.1 Principe de mesure et mode de fonctionnement	13
2.2.2 Calibration et incertitudes intrinsèques	14
2.3 SYSTÈME HIRST.....	14
2.3.1 Principe de mesure et mode de fonctionnement	14
2.3.2 Calibration et incertitudes intrinsèques	15
3 Méthodologie d'analyse des données.....	16
3.1 INTRODUCTION.....	16
3.2 DONNÉES D'ACQUISITION.....	16
3.3 DONNÉES D'ANALYSE.....	17
3.3.1 Variation temporelle et nuage de points	17
3.3.2 Calcul de l'incertitude	18
3.3.3 Analyse statistique	18
3.4 OUTILS D'ANALYSE	21
4 Inter-comparaison des résultats des capteurs.....	22
4.1 ANALYSE DES RÉSULTATS.....	22
4.1.1 Taxon Quercus.....	22
4.1.2 Taxon Fraxinus	27
4.1.3 Taxon Betula	30
4.1.4 Taxon Alnus	34
4.1.5 Taxon Corylus.....	37
4.1.6 Taxon Platanus	40
4.2 ANALYSE POUR LA PÉRIODE DU 12 MAI AU 15 SEPTEMBRE DES 3 CAPTEURS : POLLESENSE 12/95 ET SWISENS	44
4.2.1 Taxon Quercus.....	44
4.2.2 Taxon Fraxinus	44
4.2.3 Taxon Betula	45
4.2.4 Taxon Alnus	45
4.2.5 Taxon Corylus.....	46
4.2.6 Taxon Platanus	46
5 analyse des composantes principales	47

5.1	INTRODUCTION.....	47
5.2	OUTIL D'ANALYSE.....	47
5.3	STRATÉGIE D'ANALYSE.....	47
5.4	RÉSULTATS.....	48
6	Conclusion et perspectives	50

LISTE DES ILLUSTRATIONS

Figure 1	: Illustration de l'instrument et sous-systèmes d'acquisition (source Swinsens).....	8
Figure 2	: Étapes de détection et acquisitions dans la chambre de mesures.....	9
Figure 3	: Extrait de l'étape de classification avec les 24 premiers taxons.....	10
Figure 4	: Variation sur 7 jours du débit de pompage (ciblé à 40l/mn).....	12
Figure 5	: Variation au cours d'une journée (24h).....	13
Figure 6	: Illustration de l'instrument Pollensense.....	14
Figure 7	: Illustration du capteur Hirst.....	15
Figure 8	: Exemple de variation temporelle.....	17
Figure 9	: Exemple des données du RNSA.....	17
Figure 10	: Exemple de nuage de points.....	18
Figure 11	: Exemple de visualisation de l'erreur statistique pour le système Hirst.....	18
Figure 12	: Exemple de visualisation d'une matrice de corrélation.....	19
Figure 13	: Exemple de visualisation d'une BAM.....	21
Figure 14	: Données Quercus du RNSA.....	22
Figure 15	: Variation temporelle du Quercus.....	23
Figure 16	: Zoom sur la variation temporelle du Quercus du 22 avril au 1 ^{er} mai.....	23
Figure 17	: Nuage de points du Quercus.....	24
Figure 18	: Erreur statistique Quercus capteur SwisensPoleno.....	25
Figure 19	: Erreur statistique Quercus capteur Hirst.....	25
Figure 20	: Boîte à Moustache Quercus.....	26
Figure 21	: Matrice de corrélation Quercus.....	26
Figure 22	: Données Fraxinus du RNSA.....	27
Figure 23	: Variation temporelle du Fraxinus.....	27
Figure 24	: Zoom sur la variation temporelle du Fraxinus du 22 avril au 1 ^{er} mai.....	27
Figure 25	: Erreur statistique Fraxinus capteur SwisensPoleno.....	28
Figure 26	: Erreur statistique Fraxinus capteur Hirst.....	28
Figure 27	: Matrice de corrélation Fraxinus.....	29
Figure 28	: Données Betula du RNSA.....	30
Figure 29	: Variation temporelle du Betula.....	30

Figure 30 : Zoom sur la variation temporelle du Betula du 22 avril au 1 ^{er} mai.....	30
Figure 31 : Nuage de points du Betula	31
Figure 32: Erreur statistique Betula capteur SwisensPoleno	32
Figure 33 : Erreur statistique Betula capteur Hirst	32
Figure 34 : Boite à Moustache Betula	33
Figure 35 : Matrice de corrélation Betula.....	33
Figure 36 : Données de l'Alnus du RNSA	34
Figure 37 : Variation temporelle de l'Alnus.....	34
Figure 38 : Zoom sur la variation temporelle de l'Alnus du 22 avril au 1 ^{er} mai	35
Figure 39: Erreur statistique Alnus capteur Hirst	36
Figure 40 : Matrice de corrélation de l'Alnus	36
Figure 41 : Données Corylus du RNSA.....	37
Figure 42 : Variation temporelle du Corylus	37
Figure 43 : Zoom sur la variation temporelle du Corylus du 22 avril au 1 ^{er} mai.....	38
Figure 44 : Erreur statistique Corylus capteur Hirst.....	39
Figure 45 : Matrice de corrélation Corylus.....	39
Figure 46 : Données Platanus du RNSA.....	40
Figure 47 : Variation temporelle du Platanus.....	40
Figure 48 : Zoom sur la variation temporelle du Platanus du 22 avril au 1 ^{er} mai.....	41
Figure 49 : Nuage de points du Platanus	41
Figure 50 : Erreur statistique Platanus capteur SwisensPoleno	42
Figure 51 : Erreur statistique Platanus capteur Hirst	42
Figure 52 : Matrice de corrélation Platanus.....	43
Figure 53 : Variation temporelle du Quercus aout- septembre	44
Figure 54 : Variation temporelle du Betula aout- septembre.....	45
Figure 55 : Variation temporelle Alnus aout- septembre.....	45
Figure 56 : Cercle de corrélation	48

LISTE DES TABLES

Tableau 1 : Paramètres de Calibration par Taxon pour le système Poleno-23 d'AirParif.....	11
Tableau 2 : Taxons communs.....	16
Tableau 3 : Fréquence des données d'acquisition	16

NOMENCLATURE

Désignation des noms en français des taxons

Nom latin	Nom commun
Alnus	Aulne
Betula	Bouleau
Corylus	Noisetier
Fagus	Hêtres
Fraxinus	Frêne
Platanus	Platane
Poaceae	Graminée
Populus	Peuplier
Quercus	Chêne
Taxus	Ifs
Ulmus	Orme
Pinaceae	Famille : pins, épicéas, sapins, mélèzes, pruches, cèdres, douglas
Cupressus	Cyprès
Carpinus	Charme
Cupressaceae	plantes gymnospermes : 135 espèces dont cyprès, cèdre, thuya
Salix	Saule
Pinus	Pin
Urticaceae	Famille de plantes herbacées
Plantago	Plantain

1 CONTEXTE

Le laboratoire d'innovation AIRLAB d'AIRPARIF a mené une campagne d'expérimentation portant sur les mesures de pollens et ayant pour objectifs de :

- Déployer des système-capteurs de comptage de pollens permettant une qualification et une quantification en temps réel des émissions de pollens sur le territoire de l'Île-de-France ;
- Évaluer la justesse métrologique des données produites par ces capteurs par une évaluation comparative avec des mesures dites de références par la méthode « Hirst » ;
- Évaluer l'intérêt de ces nouvelles informations soit directement pour les personnes allergiques et les allergologues soit pour la prévision du risque allergique à l'échelle régionale ;
- Alimenter le développement des modèles de prévisions des concentrations en pollens dans l'air à l'échelle régionale.

Les mesures en pollens sont assurées par 3 systèmes distincts dont un en doublon.

- 1 capteur SwisensPoleno Mars
- 2 capteurs Pollensense nommés 95235 et 120015
- 1 capteur Hirst « Méthode de Référence » fourni par ATMO Grand Est

Les capteurs se situent sur le bâtiment d'AESIO mutuelle à saint Lazare (Paris 8^{ième}). L'acquisition des mesures couvre la période du 31 mars 2022 au 30 septembre 2022.

Dans ce contexte, ce rapport a pour objectif de présenter les capteurs utilisés dans le cadre de cet exercice, décrire la méthodologie utilisée pour les besoins de l'inter-comparaison des données provenant des systèmes de capteurs automatiques de pollens et présenter les résultats obtenus.

De plus, un outil d'analyse des résultats a été proposé ainsi qu'une étude préliminaire factorielle par méthode ACP.

2 CAPTEURS SUR SITE

2.1 SYSTEME SWISENSPOLENO MARS

Les éléments permettant de décrire le système SwisensPoleno sont issus des données envoyés par le fournisseur.

2.1.1 Principe de mesure et sous-systèmes associés

La solution SwisensPoleno exploite un principe de mesure couplant un prélèvement continu de volume d'air et une suite d'opérations sur le flux traversant l'équipement et visant l'acquisition de données sur les particules. Les sous-systèmes associés à ces opérations sont illustrés sur la figure suivante, extraite de la présentation de Swisens.

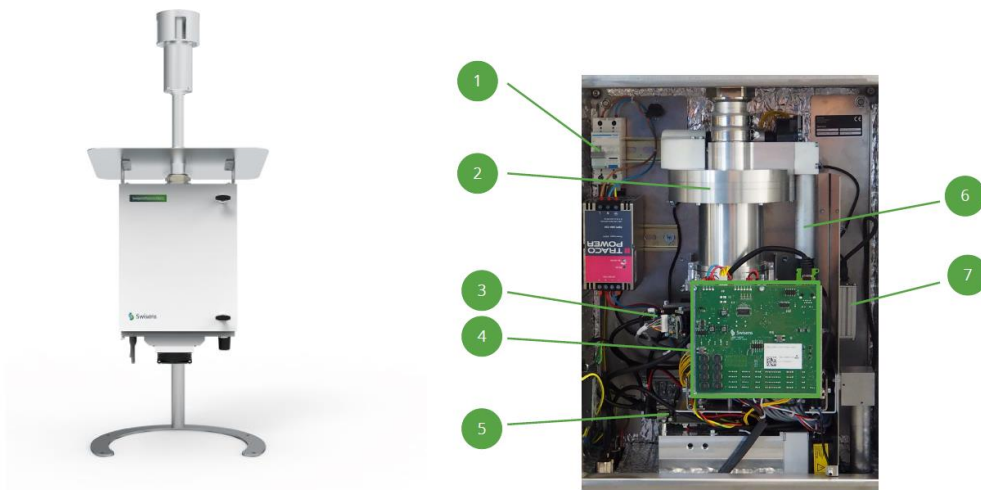


Figure 1 : Illustration de l'instrument et sous-systèmes d'acquisition (source Swisens)

Les éléments composant la partie hardware de l'équipement sont :

1. L'alimentation et système de sécurité
2. Le concentrateur de particules en aval de la tête de prélèvement Sigma2
3. La chambre de mesures (détails sur la figure suivante)
4. La carte électronique d'acquisition
5. L'UC embarquée pour la gestion/stockage et communication des données d'évènement (par particule) au data server de Swisens
6. Un bypass aéraulique
7. Le routeur de communication mobile

La technologie de prélèvement par pompage sur la tête Sigma2 est classique et calibrée à 40l/mn pour que les systèmes avals ne saturent pas mais suffisamment élevée pour que les évènements individuels soient en nombre suffisant pour la statistique sur un cumul temporel.

Le détail de la chambre de mesure est décrit ci-dessous :

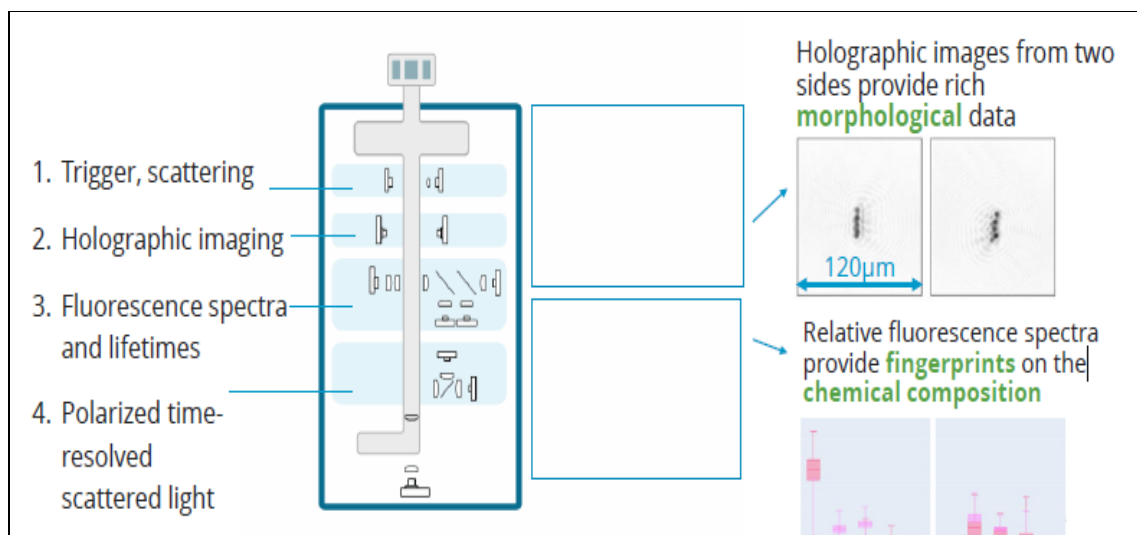


Figure 2 : Étapes de détection et acquisitions dans la chambre de mesures¹.

Les étapes séquentielles de l'acquisition de données sont :

- Le passage dans un concentrateur qui permet de focaliser le flux de particules sur une veine fluïdique restreinte.
- Une détection individuelle par diffusion qui active un imageur dual (1-Trigger).
- Une activation de l'imageur (avec 2 angles de vues à 90°) de la particule individualisée fournissant une vision « holographique » de chaque particule détectée (2-Holographic imaging).

Les systèmes collectent donc chaque évènement (ici considéré comme une particule passant dans le champ de mesure) avec un horodatage et 2 images de reconstruction morphologique. Ces évènements individuels sont ensuite exploités par un algorithme de classification de chaque particule exploitant les images (via une IA par réseau de neurones sur une modèle entraîné). L'apprentissage initial de l'IA est réalisé sur des données de pollens mono-taxon à concentrations données. Pour cette phase d'apprentissage, des pollens frais sont collectés à la source, « aérosolisés » à une concentration donnée et injectés dans l'appareil. Le modèle de classification est donc potentiellement évolutif pour l'apprentissage de nouveaux taxons.

La classification distribue les particules pour lesquelles les 2 images holographiques sont obtenues dans 14 taxons. Celles pour lesquelles une seule image est disponible ne sont pas classées.

Les taxons identifiables dans la version du modèle « MCH-Model-2022-Productive » utilisé pour le système Poleno-23 d'AirParif sont :

- | | |
|-------------|------------|
| - Alnus | - Platanus |
| - Betula | - Poaceae |
| - Carpinus | - Populus |
| - Corylus | - Quercus |
| - Cupressus | - Taxus |
| - Fagus | - Ulmus |
| - Fraxinus | |
| - Pinaceae | |

¹ Nota pour le système Poleno Mars d'Airparif, les modules de fluorescence et diffusion polarisée ne sont pas intégrés.

L'apprentissage à partir des images permet de fournir un indice de probabilité qualifiant le classement dans le bon taxon. L'image sur figure ci-dessous donne un exemple issu de classification de 2700 particules analysées le 22/04/2022 entre 00h et 12h. Seules les 24 premières particules sont illustrées.

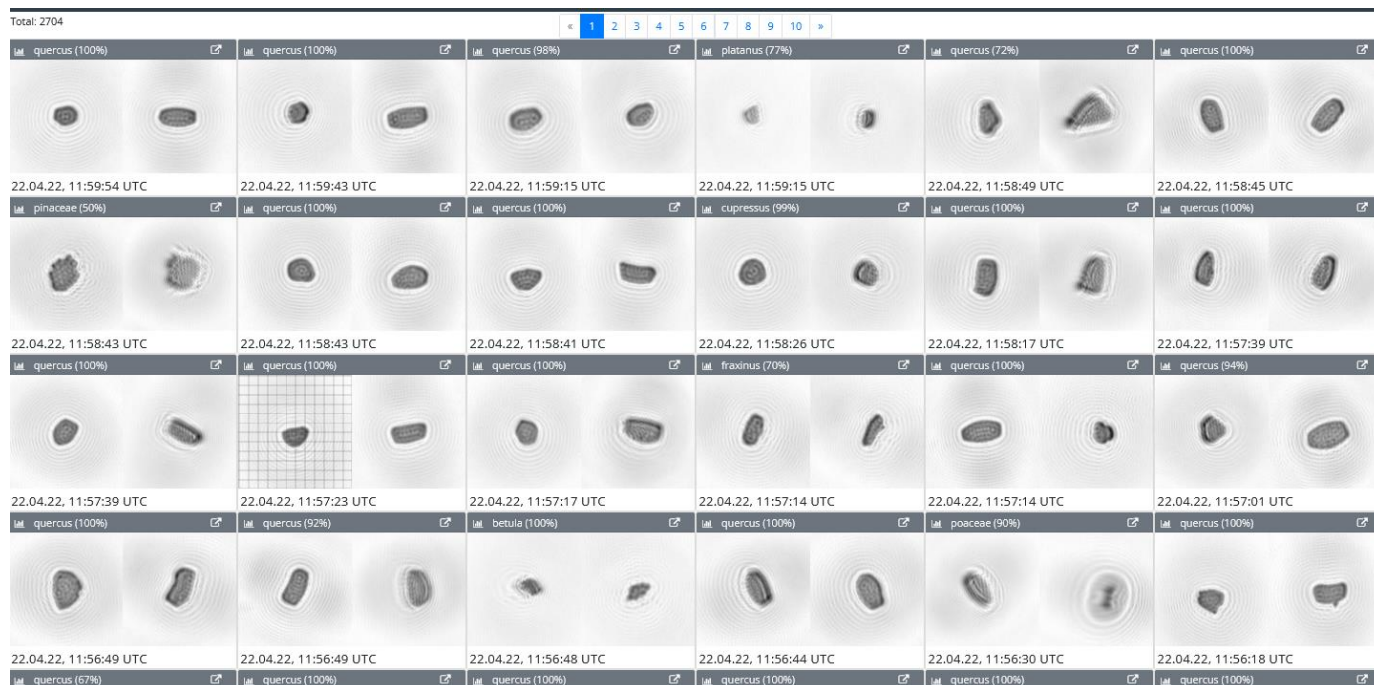


Figure 3 : Extrait de l'étape de classification avec les 24 premiers taxons

Sur chaque couple d'image, on note l'horodatage, le nom du taxon et la probabilité de succès de l'identification dudit taxon par l'algorithme de reconnaissance.

Les particules non identifiées par l'algorithme sont écartées du comptage des particules (elles ont le label « outliers » sur l'interface). Ces particules peuvent être d'origine minérale, issues de procédés de combustion, et/ou des particules de chimie hétérogène secondaire. Elles peuvent aussi être des pollens anciens (de morphologie dégradée) non attribuables.

Il faut noter que, via la tête de prélèvement et le concentrateur, les gammes de tailles de particules qui entrent dans la chambre de mesure (imageurs) vont de 2µm à 300µm. Pour un débit de 40l/mn, les gammes de concentration couvertes sans saturation vont jusqu'à 30000p/m³ pour la gamme 10µm à 300µm et jusqu'à 300000p/m³ pour les particules de 2µm.

2.1.2 Calibration et incertitudes intrinsèques (hors variabilité naturelle)

2.1.2.1 Calibration

Le principe de collecte puis de capture d'images déclenchée par diffusion et enfin d'analyse IA exploite une arrivée stochastique des particules dans la chambre de mesure. Cependant le processus global requiert une calibration des comptages pour remonter aux concentrations dans l'air par taxon.

Plusieurs facteurs peuvent induire un écart entre le nombre réel de particules pour un taxon donné entrant dans l'instrument et le nombre qualifié classé issu de l'algorithme de classement :

- Perte partielle dans le concentrateur (stockage aux parois)
- Passage hors champ (plans focaux étroits des 2 imageurs),
- Retard de déclenchement (e.g. 1/2 voir pas d'image disponible)
- Non reconnaissance par l'algorithme (certains taxons ont des formes difficiles à distinguer)

C'est la raison pour laquelle, une calibration de l'instrument est réalisée par taxon au travers :

- d'un criblage sélectionnant uniquement les sorties de l'IA avec un taux de succès élevé
- de l'application d'un paramètre de normalisation (unique et constant par taxon) déterminé sur les étapes expérimentales dédiées de calibration.

Ces 2 paramètres de calibration de l'algorithme de comptage des particules par taxon sont donnés dans le tableau suivant

Tableau 1 : Paramètres de Calibration par Taxon pour le système Poleno-23 d'AirParif

Taxon	Alnus	Betula	Carpinus	Corylus	Cupressus	Fagus	Fraxinus	Pinaceae	Platanus	Poaceae	Populus	Quercus	Taxus	Ulmus	Water
Seuil algorithmique (s.u. [0-1])	0,9	0,99	0,9	0,97	0,4	0,97	0,97	0,8	0,97	0,97	0,97	0,99	0,95	0,99	0,999
Fc Facteur de calibration (s.u.)	8	13,72	10,48	16,46	8,48	9,6	11,52	3,04	2,62	9,6	2,88	5,1	1,92	4,44	1

Les valeurs des seuils pour les critères algorithmiques sont un compromis entre le risque de mauvaise attribution et un nombre de particules criblées par l'IA suffisamment élevé. Le dernier paramètre exploité pour la reconstruction de la concentration est bien sûr lié à la durée d'acquisition T(mn) et au débit de pompage Vp [l/mn] de la tête de prélèvement.

Ainsi pour N_{ti} particules explicitement comptées par l'algorithme pour le taxon ti, la concentration en particules par m³ d'air ambiant sera : C_{ti} [p/m³] = 1000*F_{ci}*N_{ti}/(Vp*T) (le facteur 1000 opérant la conversion litre vs m³)

Par exemple, pour le chêne 1000 particules sont comptées dans un intervalle cumulé de 60mn et pour un débit de pompage de 40l/mn, la concentration sera : C[p/m³] = 1000*5.1*1000/40/60= 2125 p/m³

2.1.2.2 Incertitudes

Les données extraites du serveur de Swisens sont déjà converties pour fournir la concentration ambiante. Cependant pour l'estimation du bruit de mesure intrinsèque à l'équipement, il faut se baser sur le nombre de particules réellement comptées sur la période de cumul, hors calibration.

Les erreurs possibles sur la valeur de concentration en sorties peuvent être liées à une erreur d'estimation de F_{ci}, de Vp, de T et de N_{ti}.

- Paramètres F_{ci} : Il n'est pas possible d'accéder aux erreurs d'estimation (biais ou écart type relatif) des coefficients de calibration par taxon. Ils sont issus d'une procédure spécifique sur protocoles et essais contrôlés réalisés par Swisens avant installation.

- Paramètres T durée de cumul d'échantillonnage : L'horodatage des images est à la seconde près. Pour les besoins d'interprétation, les cumuls horaires et, au mieux de 10mn, seront utilisés. L'incertitude relative sur le paramètre T, est donc inférieure à $1s/600s$ pour un cumul sur 10mn soit 0,166% (et 10 fois moins pour un séquençage horaire). On peut donc raisonnablement la négliger devant l'incertitude relative de l'échantillonnage et/ou la variabilité naturelle sur de telles durées.
- Paramètre vitesse de pompage V_p : Le contrôleur de l'équipement cible des débits volumiques à 40l/mn. Lors d'opération de maintenance automatique journalière autour de 01h UTC, un débit de nettoyage élevé est appliqué pendant moins d'1mn. Les poussières potentiellement accumulées dans les veines fluidiques sont ainsi éliminées. Les données associées à ces périodes de maintenance ne sont pas archivées ni extractibles. Un exemple sur 7 jours consécutifs est fourni ci-dessous (Figure 4) . Les pics correspondent au flushing des veines fluidiques (passage à >120 l/mn pendant moins d'une minute). Les variations dans la journée (figure 5) montrent une amplitude de l'ordre de 1,2l/mn et un bruit de mesure haute fréquence de débit d'amplitude 0,1l/mn.

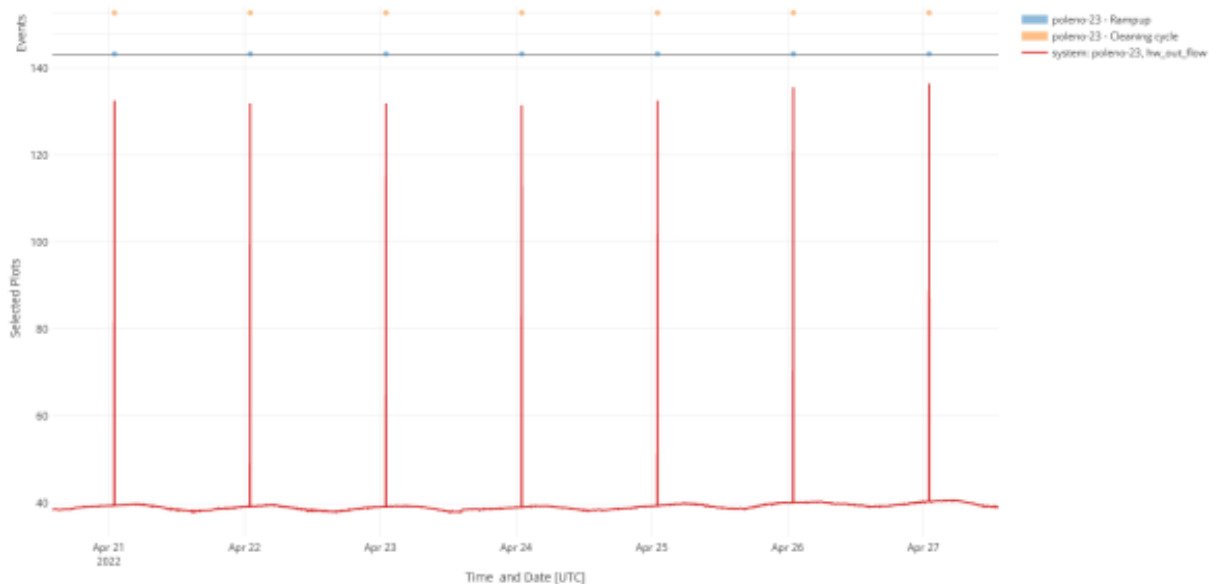


Figure 4 : Variation sur 7 jours du débit de pompage (ciblé à 40l/mn)

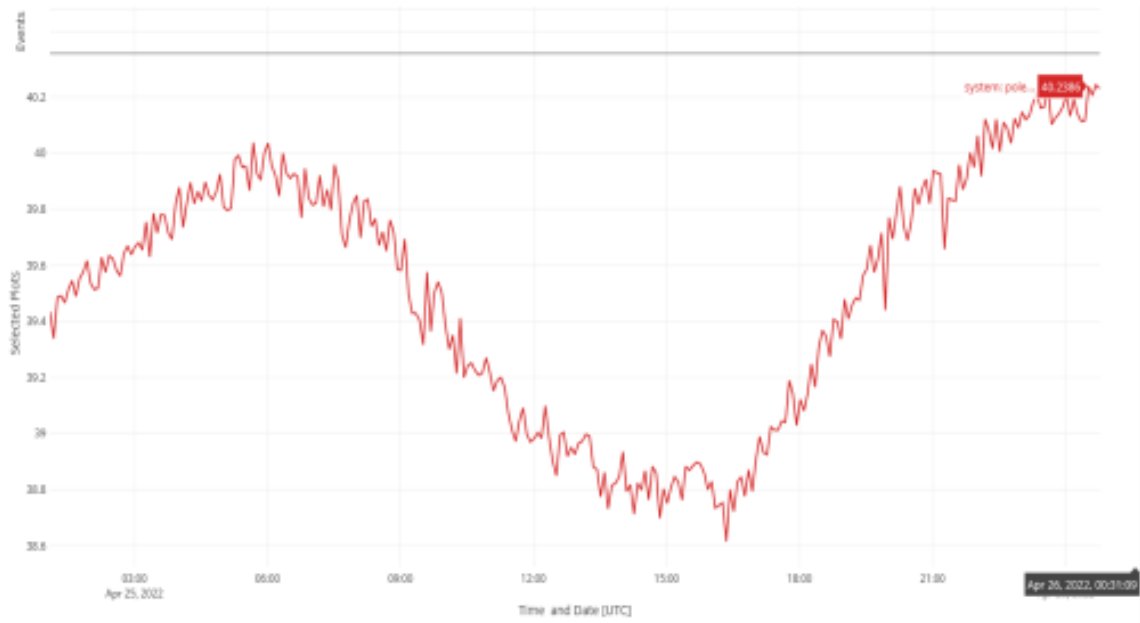


Figure 5 : Variation au cours d'une journée (24h)

Une faible dérive peut être observée dans la durée (pour 24h) sur le débit volumique. Elle présente une amplitude de l'ordre de 1,2 l/mn. La construction de la concentration en sortie d'analyse en tient compte. Le bruit de mesure sur le contrôleur (fluctuations haute fréquence sur le graphe de droite de la figure est de l'ordre de 0,1l/mn en amplitude min/max. soit 0,25% en valeur relative. Ce bruit de mesure sur le débit de pompage est négligeable.

- Erreur statistique : l'hypothèse raisonnable pour définir l'incertitude de mesure que le processus sous-jacent est de type Poisson, l'écart-type de mesure pour N_{ti} particules comptée pour le taxon i est en donc $\sqrt{N_{ti}}$. Cet écart-type renormalisé par les coefficients de calibration fournit une barre d'erreur sur les courbes temporelles des concentrations pour chacun des taxons, et produites dans le cadre de l'inter-comparaison.
Par exemple sur la valeur de

2.2 SYSTEME POLLENSENSE

Les éléments permettant de décrire le système Pollensense sont issus des informations fournies par le fournisseur ainsi que celles disponibles sur leur plateforme. Il y a très peu d'informations documentées sur le principe de fonctionnement et les méthodes d'analyses associées.

2.2.1 Principe de mesure et mode de fonctionnement

La solution Pollensense exploite un principe de mesure couplant un prélèvement continu d'un volume d'air avec un débit de 14l/min. Les pollens sont ensuite fixés sur une bande adhésive .

La solution présente un système de reconstruction 3D de l'image qui est ensuite traitée par une intelligence artificielle qui classe automatiquement les espèces.



Figure 6 : Illustration de l'instrument Pollensense

Les taxons identifiables pour le système Pollensense sont :

- | | |
|------------|----------------|
| - Alnus | - Cupressaceae |
| - Betula | - Salix |
| - Corylus | - Urticaceae |
| - Fagus | - Grass |
| - Fraxinus | - Pollen |
| - Platanus | - Weed.Shrub |
| - Poaceae | - Plantago |
| - Populus | - Tree |
| - Quercus | - Pinus |

2.2.2 Calibration et incertitudes intrinsèques

Ce système utilise également une calibration entre le nombre réel de particules pour un taxon donné entrant dans l'instrument et le nombre qualifié issu de l'algorithme de classement. Le détail sur le principe de calibration ainsi que les facteurs associés ne sont pas disponibles. L'utilisateur n'a pas accès à la variation du débit de pompage. Il n'est donc pas possible de calculer l'incertitude sur la mesure physique.

2.3 SYSTEME HIRST

2.3.1 Principe de mesure et mode de fonctionnement

L'air est aspiré au moyen d'une pompe, selon un débit stable et régulier de 10 litres d'air par minute, ce qui correspond à une respiration humaine moyenne.

L'air entre dans l'appareil au moyen d'une buse d'aspiration. Les grains de pollen qu'il contient sont alors projetés sur une bande transparente adhésive, comparable à une bande de scotch, entraînée par un tambour.

La rotation de ce tambour est assurée par un mécanisme d'horlogerie, à raison de 2 mm par heure. Le tambour effectue un tour complet en une semaine. C'est grâce à ce système que l'analyse temporelle des prélèvements est possible.

Pour permettre au capteur d'aspirer les grains de pollens de manière optimale, la buse d'aspiration est constamment orientée face aux vents dominants.

Chaque semaine, la bande sur laquelle sont récoltés les grains de pollens est collectée. Son analyse microscopique permet d'identifier et de compter les espèces végétales présentes sur les prélèvements et à établir une prévision pour la semaine suivante.

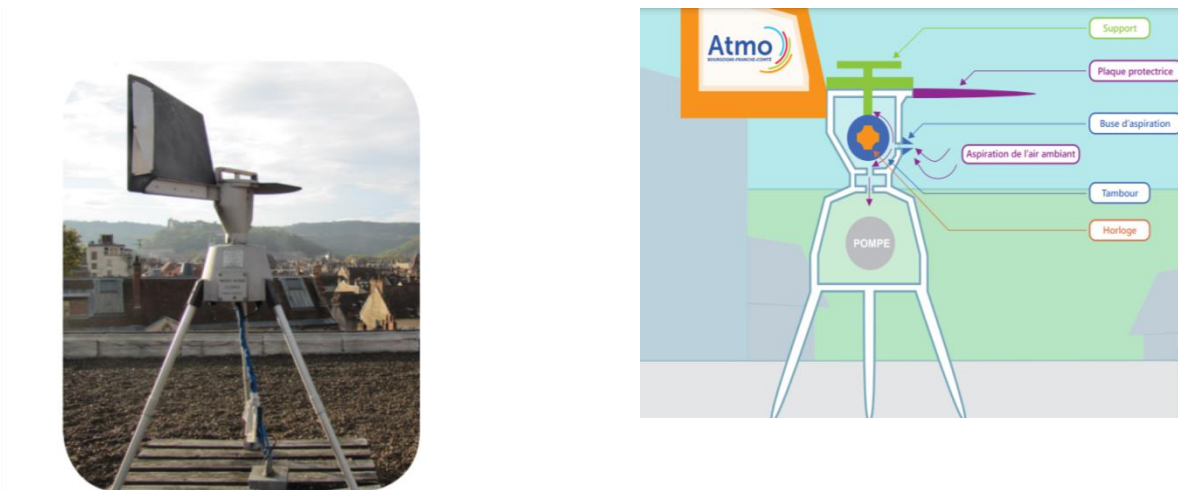


Figure 7: Illustration du capteur Hirst

Le système Hirst permet d'identifier une centaine d'espèces.

2.3.2 Calibration et incertitudes intrinsèques

Ce système utilise également une calibration entre le nombre réel de particules pour un taxon donné entrant dans l'instrument et le nombre qualifié issu de l'algorithme de classement. Pour cet instrument il est utilisé un facteur de calibration de 0.58 qui est identique pour chacun des taxons.

Concernant les variations de débit de pompage elle est observée lors du changement de la bande adhésive. On note une variation de 0.5 à 1,5 L/mn ce qui est négligeable.

3 METHODOLOGIE D'ANALYSE DES DONNEES

3.1 INTRODUCTION

L'analyse des résultats a été établie pour l'ensemble des 4 capteurs :

- SwisensPoleno Mars
- Pollensense n° 95235
- Pollensense n° 120015
- Hirst

Les espèces mesurées par chacun des 3 systèmes étant différentes, pour les besoins de cette comparaison seuls les taxons communs entre les systèmes sont retenus. Le tableau suivant illustre les taxons communs.

Tableau 2 : Taxons communs

Betula	Corylus	Quercus	Fraxinus	Platanus	Alnus
Bouleau	Noisetier	Chêne	Frêne	Platane	Aulne

Les analyses portent sur :

- La variation temporelle des concentrations en taxons sur l'ensemble des 4 capteurs
- Le nuage de points et les courbes de tendance en référence au système Hirst
- Le calcul de l'erreur statistique pour les systèmes Swisens et Hirst
- L'analyse statistique sous forme de boîte à moustache
- L'analyse statistique sous forme matrice de corrélation

3.2 DONNEES D'ACQUISITION

Les données d'acquisition ne sont pas uniformes selon les systèmes. Le tableau suivant illustre les pas de temps d'acquisition pour chacun des capteurs.

Tableau 3 : Fréquence des données d'acquisition

Système	Pas de temps d'acquisition des données								
Swisens	10 min	15 min	30 min	1h	2h	3h	6h	12h	24h
Pollensense	-	-	-	1h	-	-	-	-	24h
Hirst	-	-	-	-	2h	-	-	-	-

Afin de mettre en place les éléments de comparaison, il est impératif de travailler sur un pas de temps commun. Le système Hirst ne fournissant que des données de manière bihoraire, le pas de temps d'acquisition de 2 heures est retenu.

Le système Pollensense ne permet pas d'extraire les données toutes les 2 heures. Il est donc nécessaire d'établir un premier travail de mise en forme des données afin de retranscrire les concentrations toutes les 2 heures sur la base des données horaires.

3.3 DONNEES D'ANALYSE

3.3.1 Variation temporelle et nuage de points

Les premières analyses portent sur les variations temporelles ainsi que sur les nuages de points.

Les variations temporelles fournissent les concentrations en $\mu\text{g}/\text{m}^3$ pour la période d'acquisition donnée.

Afin d'avoir une information plus précise sur la validité de la présence ou non du taxon, ces données sont comparées au calendrier pollinique issue du RNSA pour l'année 2022 (<https://pollens.fr/les-risques/risques-par-ville/40/8/2022>).

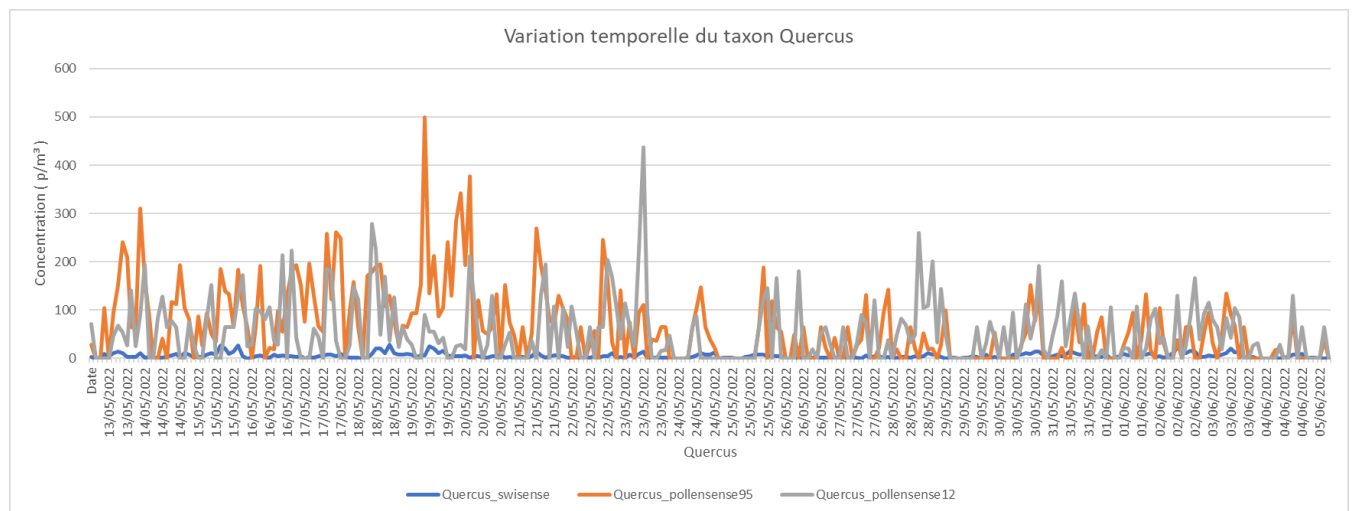


Figure 8 : Exemple de variation temporelle

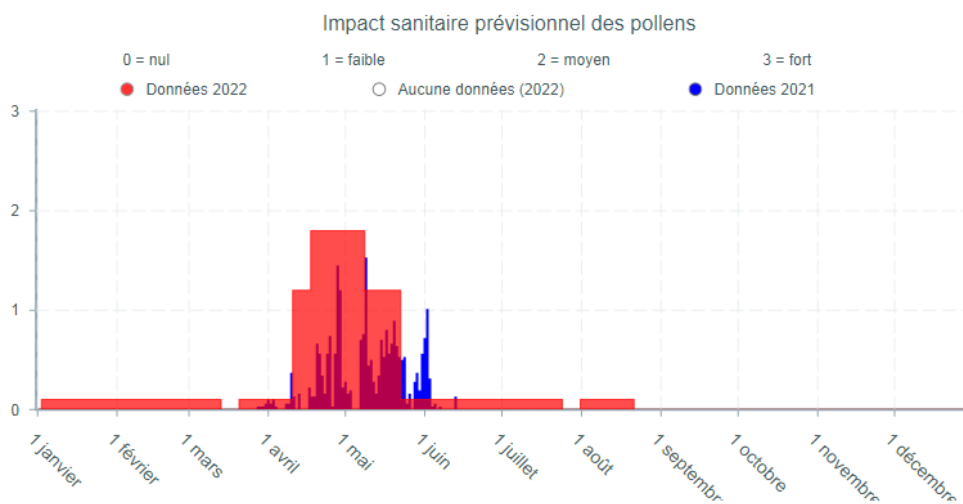


Figure 9 : Exemple des données du RNSA

Le nuage de points représente les valeurs détectées par chaque système par rapport aux valeurs de référence Hirst.

La projection de chaque de chaque point sur l'axe horizontal donne la valeur de Hirst et la projection sur l'axe vertical donne la valeur pour les trois autres systèmes.

Cette représentation permet de réaliser une comparaison avec Hirst en fonction des concentrations mesurées et d'évaluer le degré de déviation de chaque système par rapport au système de référence Hirst.

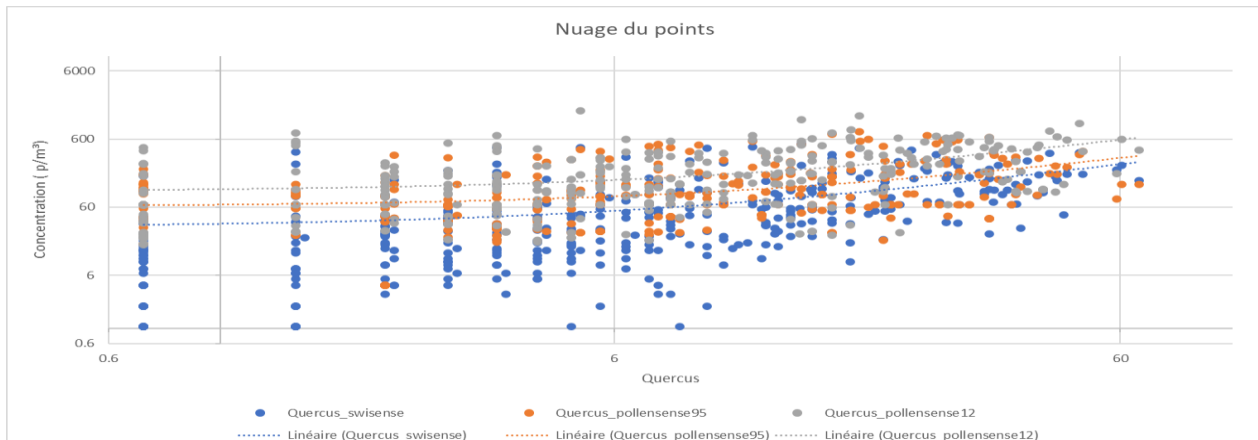


Figure 10 : Exemple de nuage de points

3.3.2 Calcul de l'incertitude

Le calcul de l'incertitude est focalisé sur le paramètre intrinsèque de la calibration du nombre de particule. Au vu des données constructeur, l'erreur statistique ne peut être calculée que pour le système Hirst et SwisensPoleno.

L'erreur est calculée à partir de la relation suivante : $\text{Erreur} = \sqrt{N}$

N : le nombre de particules détecté par les capteurs

Les courbes représentent :

- La courbe d'erreur de surestimation = $N + \sqrt{N}$
- La courbe d'erreur de sous-estimation = $N - \sqrt{N}$

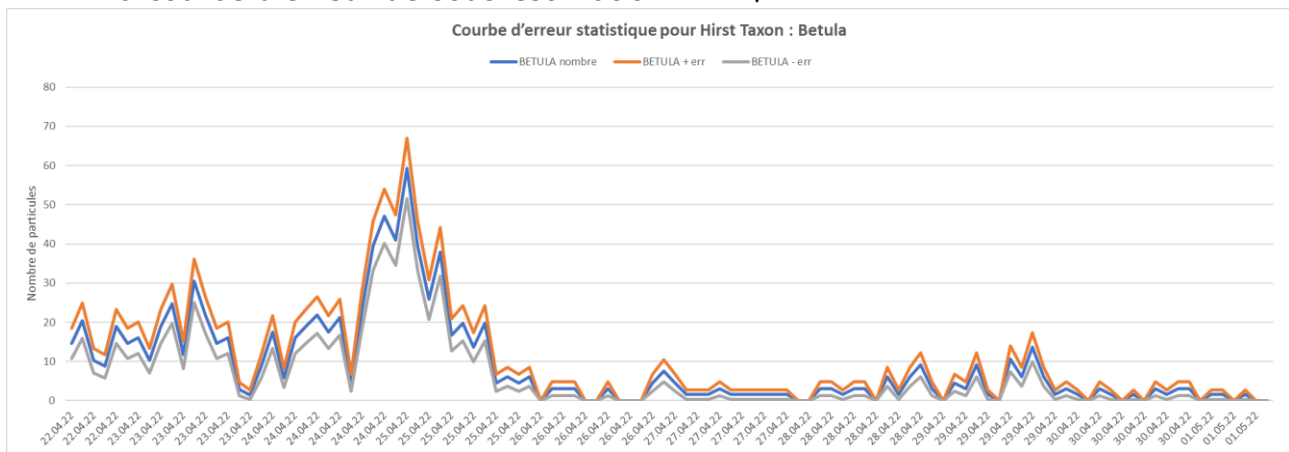


Figure 11 : Exemple de visualisation de l'erreur statistique pour le système Hirst

3.3.3 Analyse statistique

L'analyse statistique repose sur la mise en place de boîte à moustache et de matrice de corrélation.

La matrice de corrélation est une matrice des coefficients de corrélation statistiques calculés sur plusieurs variables prises deux à deux. Dans notre cas la

matrice de corrélation représente les différents coefficients de corrélation (Coefficient de Pearson) pour une combinaison d'un binôme de capteur. Les coefficients de corrélation varient entre -1 à + 1, les valeurs positives représentant le même sens des tendances et les valeurs négative un sens inversé des tendances.

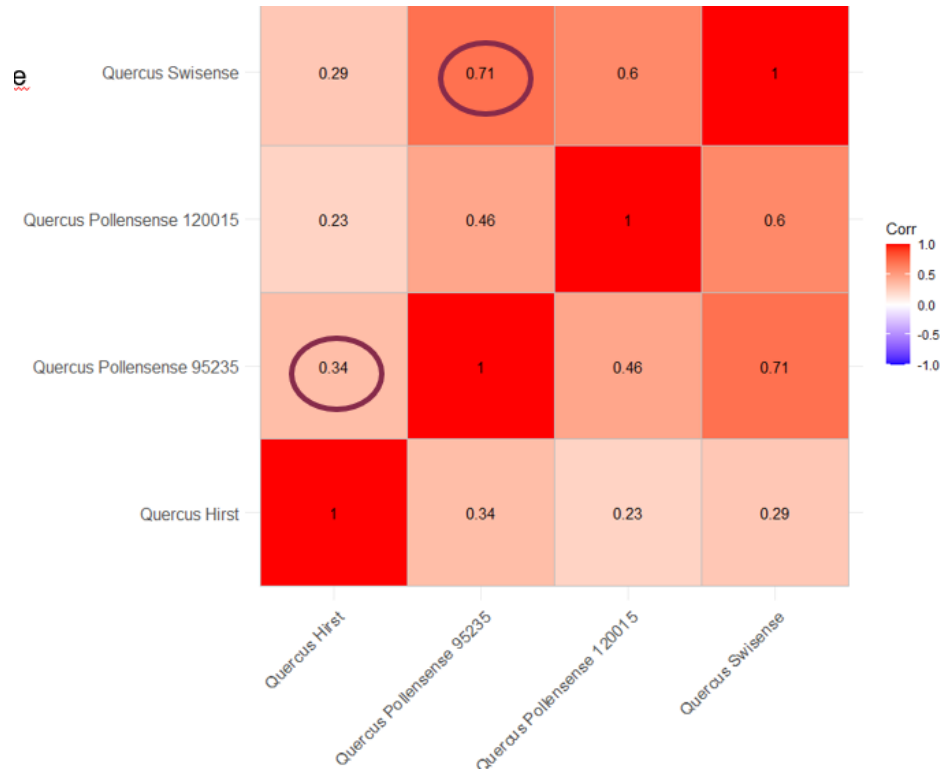
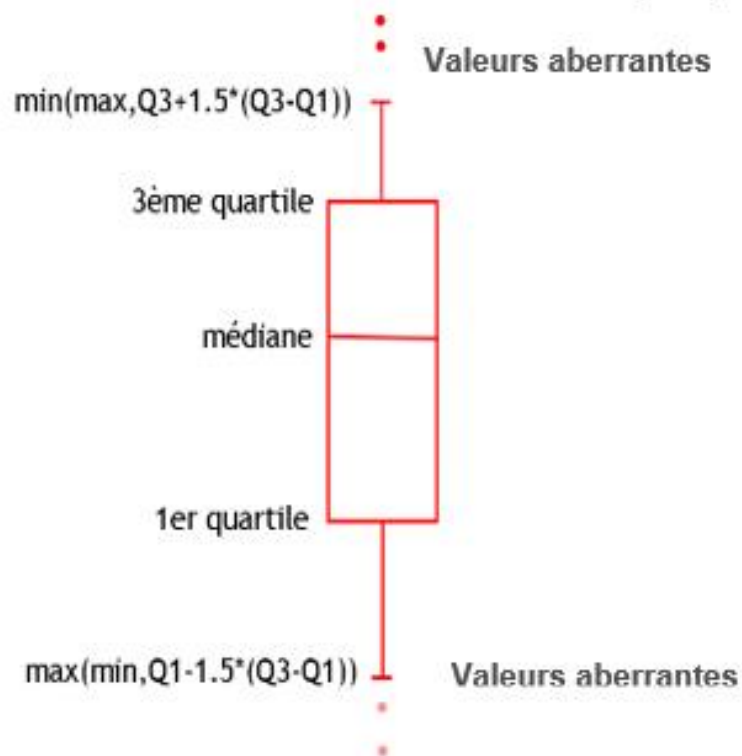


Figure 12 : Exemple de visualisation d'une matrice de corrélation

Le second outil statistique est représenté par les boîtes à moustache. Il s'agit d'un moyen rapide de figurer le profil essentiel d'une série statistique quantitative. En effet, ce type de visualisation résume en seulement quelques indicateurs de position le caractère étudié (médiane, quartiles, minimum, maximum).



La relation étudiée dans le cas de nos analyses statistiques est :

$$\text{Différence Relative} = \text{abs} ((\text{capteur } x - \text{capteur } y) / (\text{capteur } x + \text{capteur } y)) * 2$$

Les boîtes à moustache visualisent la différence relative entre les systèmes:

- Pollensense 95235 et Hirst
- Pollensense 120015 et Hirst
- SwisensPoleno et Hirst
- Pollensense 95235 et Pollensense 120015
- Pollensense 95235 et SwisensPoleno
- Pollensense 120015 et SwisensPoleno

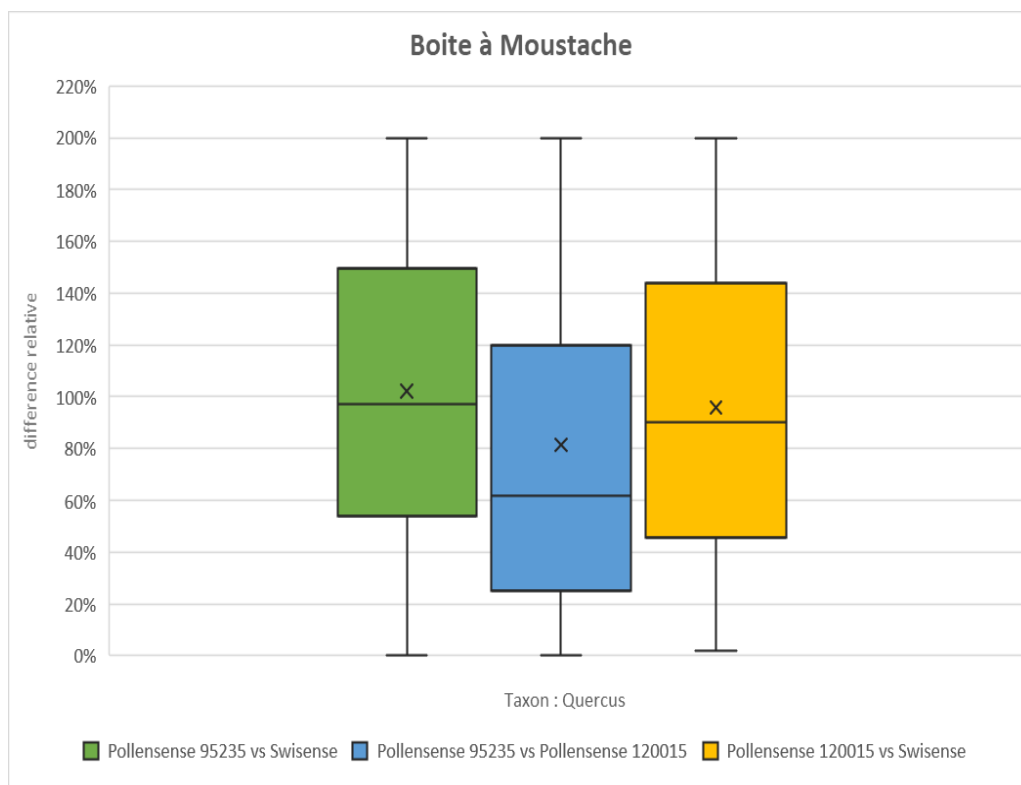


Figure 13 : Exemple de visualisation d'une BAM

3.4 OUTILS D'ANALYSE

Un outil permettant de comparer les données de chacun des systèmes est fourni au client. Cet outil est présenté sous format Excel nommé « **Pollen-outil statistique.xlsx** ». Il est accompagné d'un guide utilisateur « **Fluidyn-Guide-utilisateur_analyse-Pollen_1.0.doc** ».

Cet outil permet de réaliser la comparaison entre les quatre systèmes de comptage en pollens qui sont :

- SwisensPoleno
- Pollensense 95
- Pollensense 12
- Hirst

Le principe de comparaison repose sur l'analyse de chacun des systèmes avec celui de Hirst qui est supposé comme celui de référence pour ces travaux.

La comparaison se fait pour un ensemble de 6 taxons qui représentent les espèces communes à l'ensemble des 4 systèmes.

Les graphiques automatiques fournies par l'outil sont :

- Les graphiques des séries temporelles
- Les nuages de points
- Les erreurs statistiques
- Les boîtes à moustache
- Les matrices de corrélations

4 INTER-COMPARAISON DES RESULTATS DES CAPTEURS

Des analyses d'inter-comparaison ont été menées au fur et à mesure de l'acquisition des données de chacun des 4 capteurs.

Pour rappel les données d'expérimentales sont disponibles de fin mars à septembre 2022. Certains tambours ayant été perdus, le capteur Hirst ne fourni pas de données pour la période 27 juillet au 4 aout.

Le détail des résultats des analyses sont présentés dans les documents envoyés au client comprenant :

- **Analyse 1** : Données d'acquisition du 31 mars au 14 avril 2022 pour l'ensemble des 4 systèmes
- **Analyse 2** : Données d'acquisition du 22 avril au 1 mai 2022 pour l'ensemble des 4 systèmes
- **Analyse 3** : Données d'acquisition du 31 mars au 12 mai 2022 pour l'ensemble des 4 systèmes
- **Analyse 4** : Données d'acquisition du 12 mai au 15 septembre 2022 pour les systèmes Pollensense et Swisens
- **Analyse 5** : Données d'acquisition du 12 mai au 27 juillet 2022 et du 4 aout à septembre 2022 pour l'ensemble des 4 systèmes

4.1 ANALYSE DES RESULTATS

Ce chapitre résume les résultats obtenus pour l'ensemble des 4 capteurs pour la période du 31 mars au 27 juillet 2022. Il s'agit donc d'une compilation des données reçues lors des analyses 1 à 5.

4.1.1 Taxon Quercus

4.1.1.1 Variation temporelle

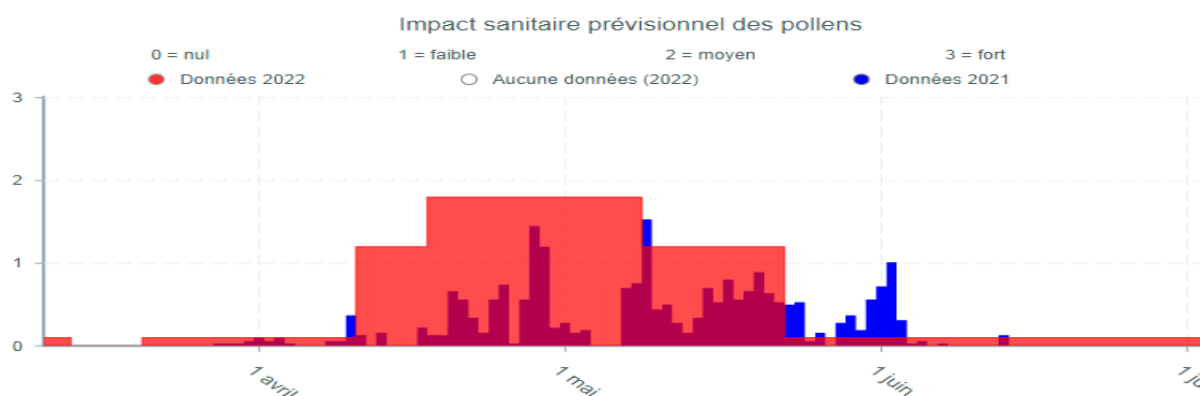


Figure 14 : Données Quercus du RNSA

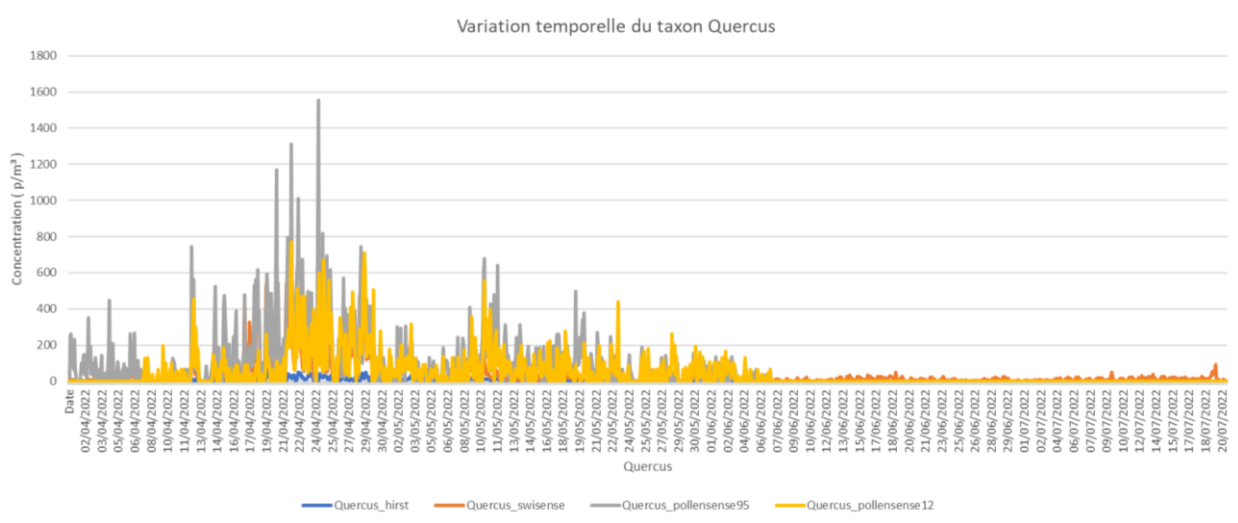


Figure 15 : Variation temporelle du Quercus

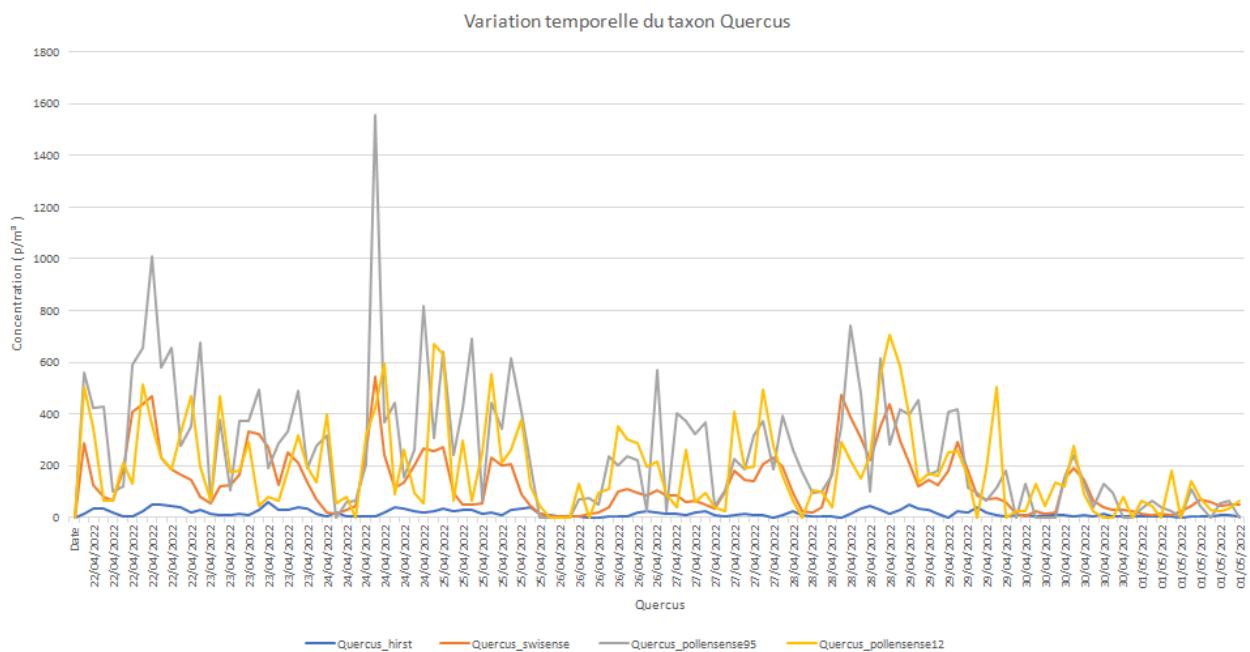


Figure 16 : Zoom sur la variation temporelle du Quercus du 22 avril au 1^{er} mai

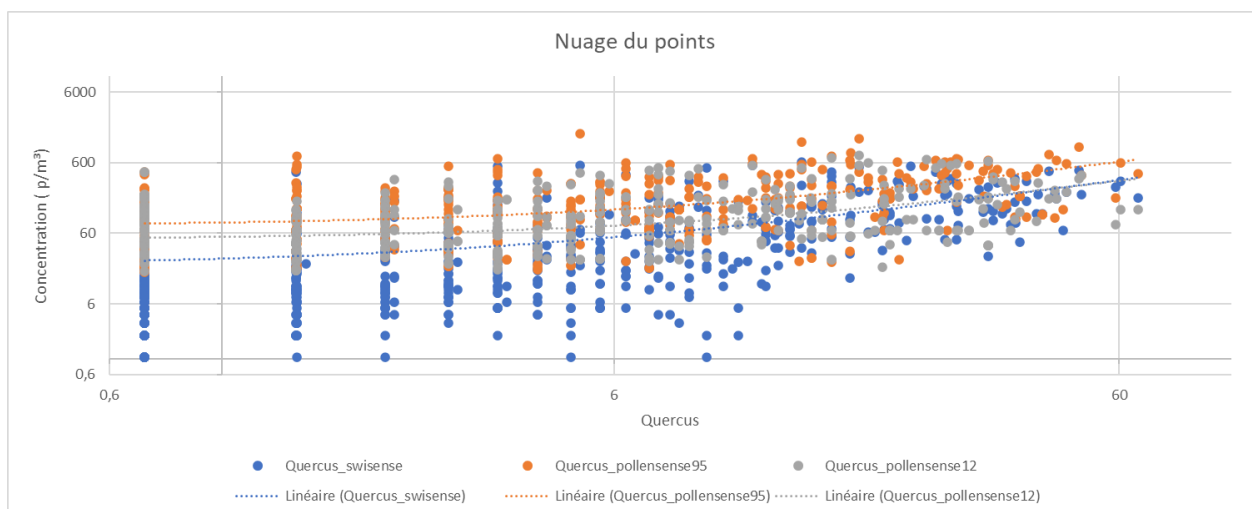


Figure 17 : Nuage de points du Quercus

D'après les données du RNSA, la saison pollinique du Quercus couvre la période allant de mi-avril à mi-mai. L'expérimentation a donc eu lieu pendant la période de floraison du Quercus.

Concernant les résultats de mesure, de manière générale, on note une différence d'intensité des concentrations. Les concentrations pour le capteur Hirst sont beaucoup plus faibles en comparaison aux autres capteurs. Les capteurs Pollensense sont ceux qui enregistrent des concentrations les plus élevées.

Le graphique zoomé pour les journées du 22 avril au 1^{er} mai montre, pour l'ensemble des capteurs, une variation cyclique durant la journée qui est certainement lié à l'aspect biologique et phénologique du pollen.

Le nuage de points reflète également une surestimation des concentrations par rapport au système Hirst. On note une meilleure corrélation pour des concentrations élevées, cela est peut être lié aux facteurs de conversion appliqués qui sont éventuellement plus adaptés à des concentrations élevées en pollen.

4.1.1.2 Erreur statistique

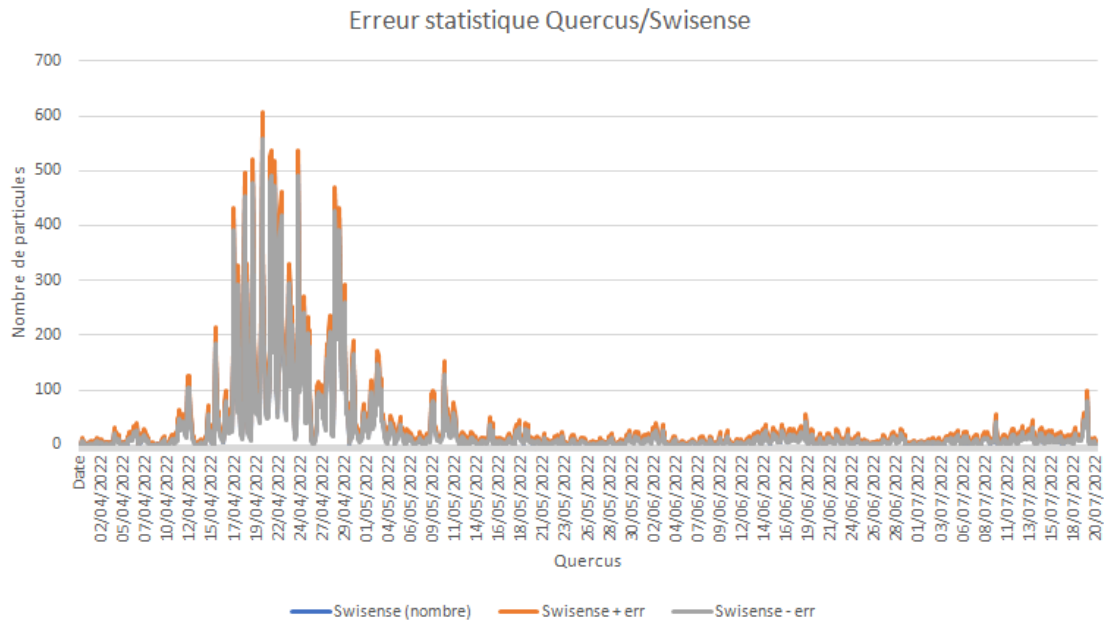


Figure 18 : Erreur statistique Quercus capteur SwisensPoleno

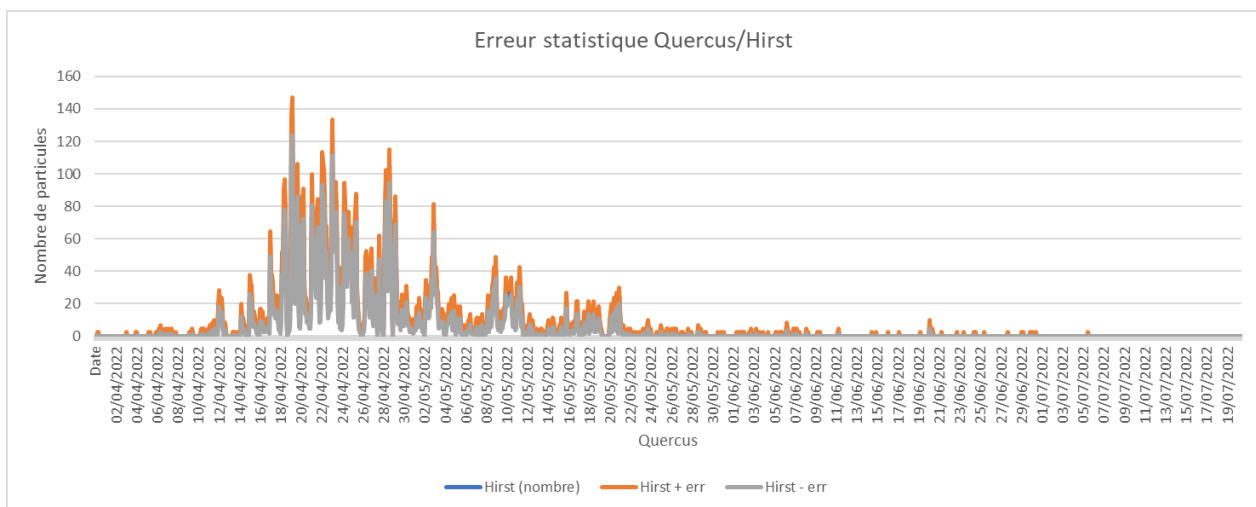


Figure 19 : Erreur statistique Quercus capteur Hirst

Les graphiques représentant les erreurs statistiques montrent un bruit sur la mesure qui est négligeable. La variabilité observée est liée à la variabilité naturelle de l'espèce.

4.1.1.3 Données statistiques

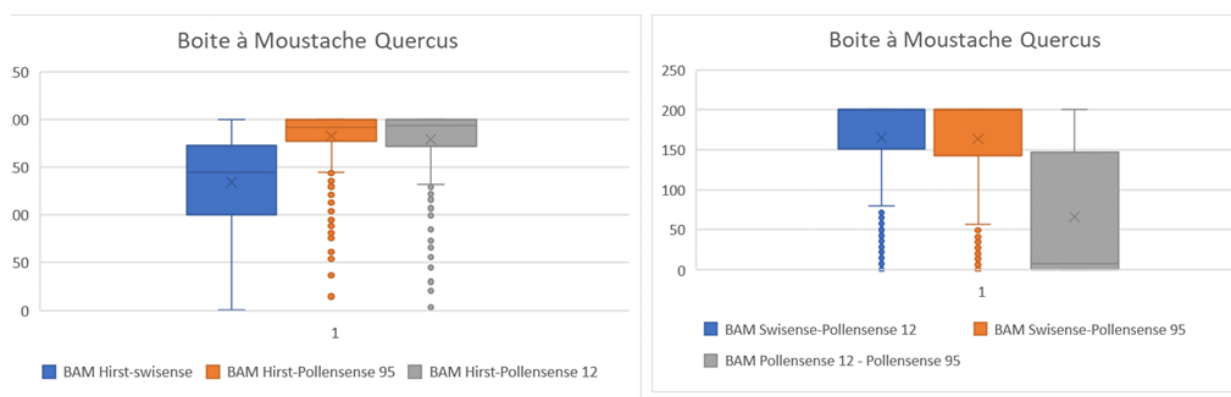


Figure 20 : Boite à Moustache Quercus

L'analyse des boites à moustache permet de détacher les conclusions suivantes :

- Concernant les comparaisons avec le capteur Hirst, pour 50% des données, la différence relative de Swisens est inférieure à celle des capteurs Pollensense.
- Concernant les comparaisons Swisens/Pollensense, pour 50% des données, la différence relative entre les capteurs Pollensense 120015 et Pollensense 95235 est la plus basse.

	Matrice de Corrélation Quercus			
	Hirst	Swisense	Pollensense 95	Pollensense 12
Hirst	1,00	0,57	0,51	0,40
Swisense	0,57	1,00	0,71	0,54
Pollensense 95	0,51	0,71	1,00	0,53
Pollensense 12	0,40	0,54	0,53	1,00

Figure 21 : Matrice de corrélation Quercus

D'après la matrice de corrélation, le couple de capteurs le mieux corrélé est le couple Swisens/ Polensense95 avec un facteur de corrélation de 0.71.

4.1.2 Taxon Fraxinus

4.1.2.1 Variation temporelle

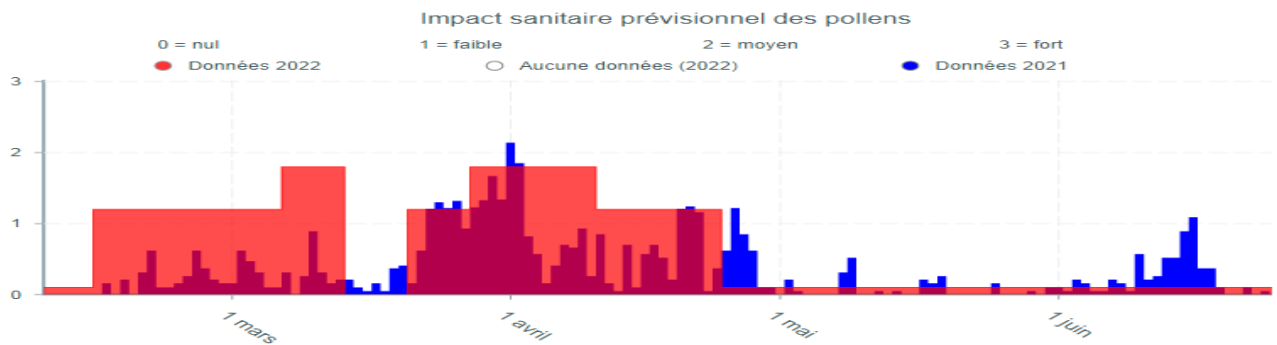


Figure 22 : Données Fraxinus du RNSA

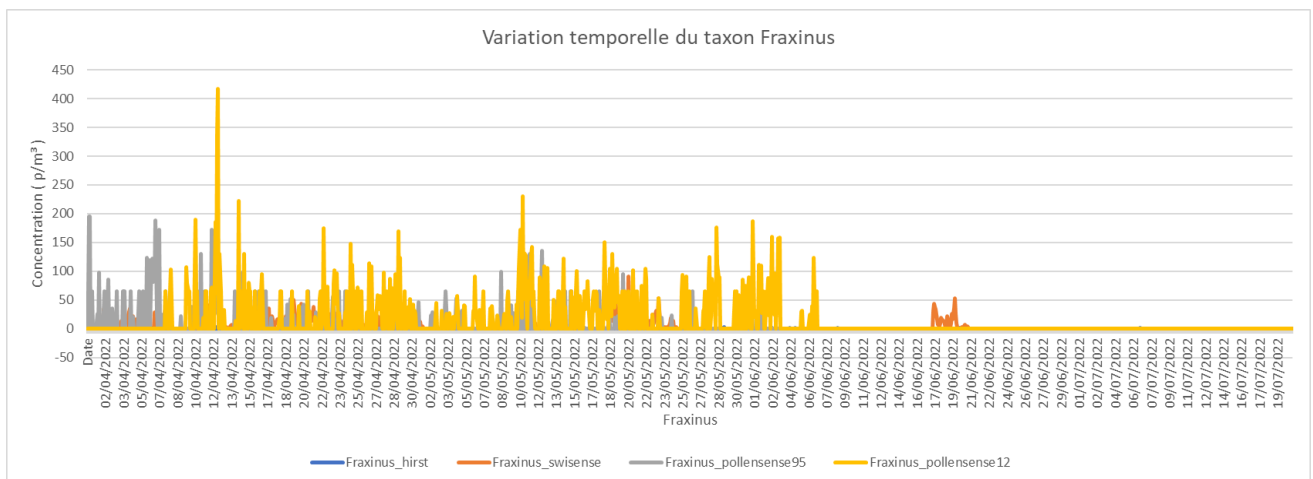


Figure 23 : Variation temporelle du Fraxinus

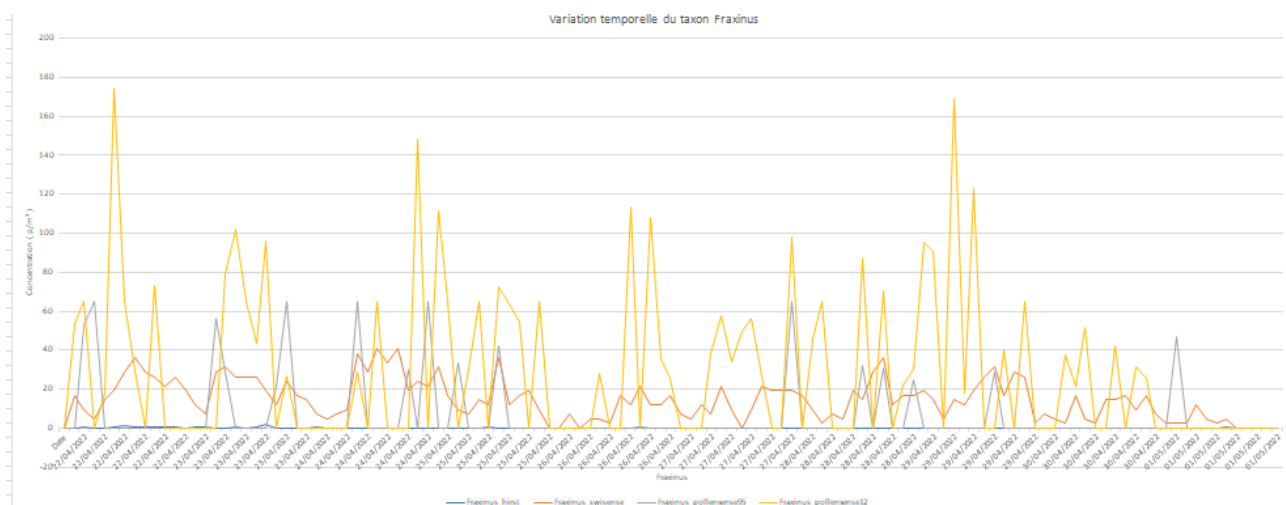


Figure 24 : Zoom sur la variation temporelle du Fraxinus du 22 avril au 1^{er} mai

D'après les données du RNSA, la saison pollinique du Fraxinus se termine mi-avril. L'expérimentation aurait donc débuté en fin de saison pour ce taxon.

Concernant les résultats de mesure, on note des concentrations nulles ou très faibles pour les systèmes Swisens et Hirst. Seuls les capteurs Pollensense enregistrent des concentrations élevées.

Le nuage de points n'est pas applicable pour ce cas au vu du faible nombre de valeurs non nulles.

4.1.2.2 Erreur statistique

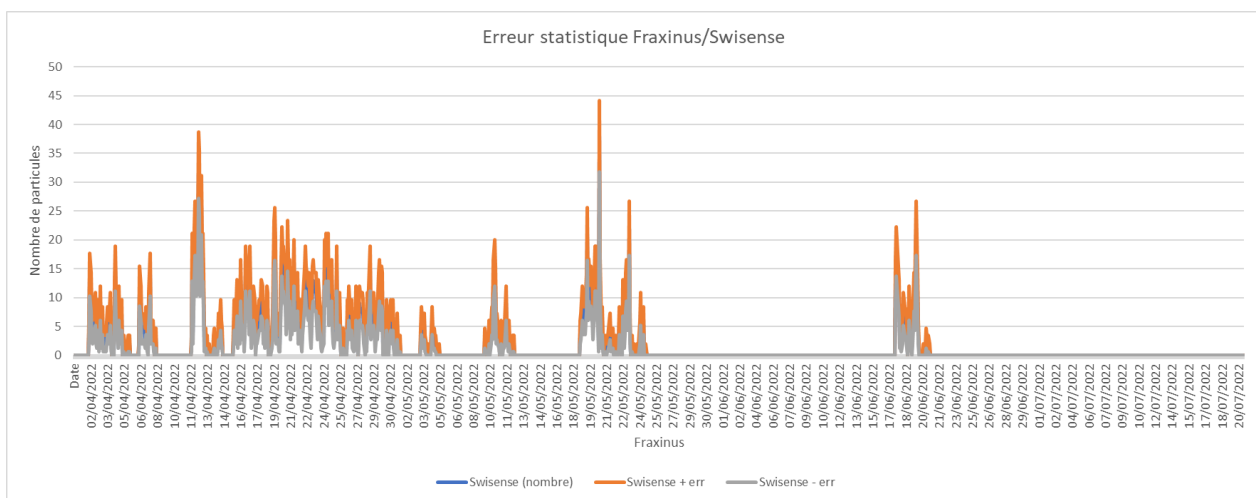


Figure 25 : Erreur statistique Fraxinus capteur SwisensPoleno

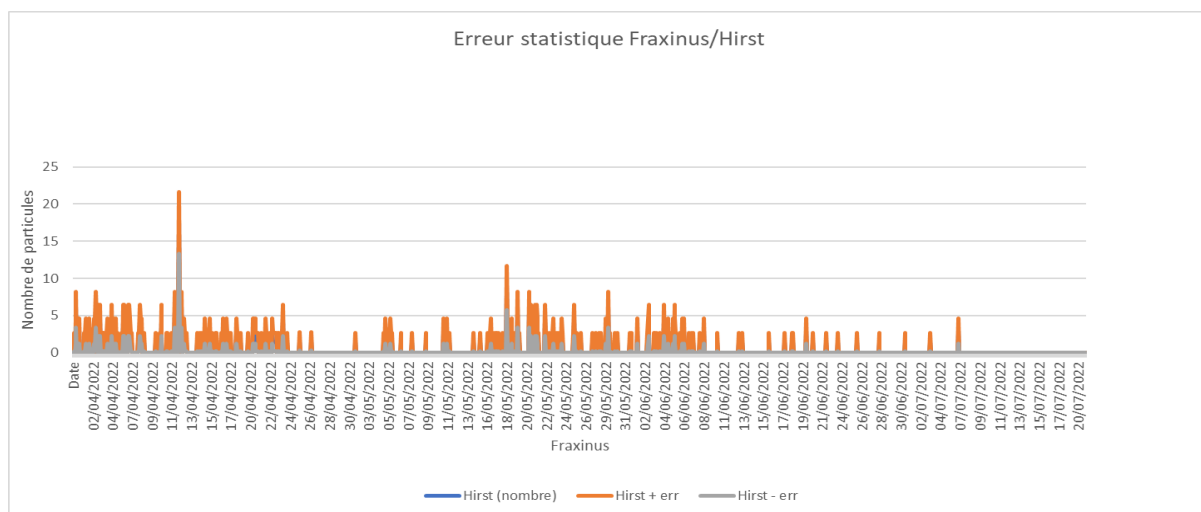


Figure 26 : Erreur statistique Fraxinus capteur Hirst

Les graphiques représentant les erreurs statistiques montrent un bruit sur la mesure qui est négligeable.

4.1.2.3 Données statistiques

Au vu du nombre de valeurs nulles sur deux des 4 systèmes, les boîtes à moustache ne sont pas applicables. Concernant la matrice de corrélation, les facteurs de corrélation sont très faibles voir même négatifs. Le couple de capteurs le mieux corrélé est le couple Swisens/ Pollensense 12 avec un facteur de corrélation de 0.12.

	Matrice de Corrélation Fraxinus			
	Hirst	Swisense	Pollensense 95	Pollensense 12
Hirst	1,00	0,04	0,03	-0,16
Swisense	0,04	1,00	0,09	0,12
Pollensense 95	0,03	0,09	1,00	-0,04
Pollensense 12	-0,16	0,12	-0,04	1,00

Figure 27 : Matrice de corrélation Fraxinus

4.1.3 Taxon Betula

4.1.3.1 Variation temporelle

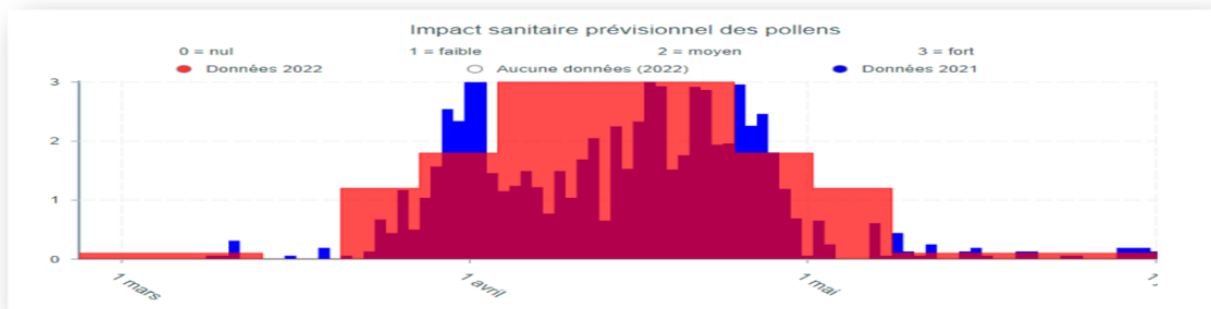


Figure 28 : Données Betula du RNSA

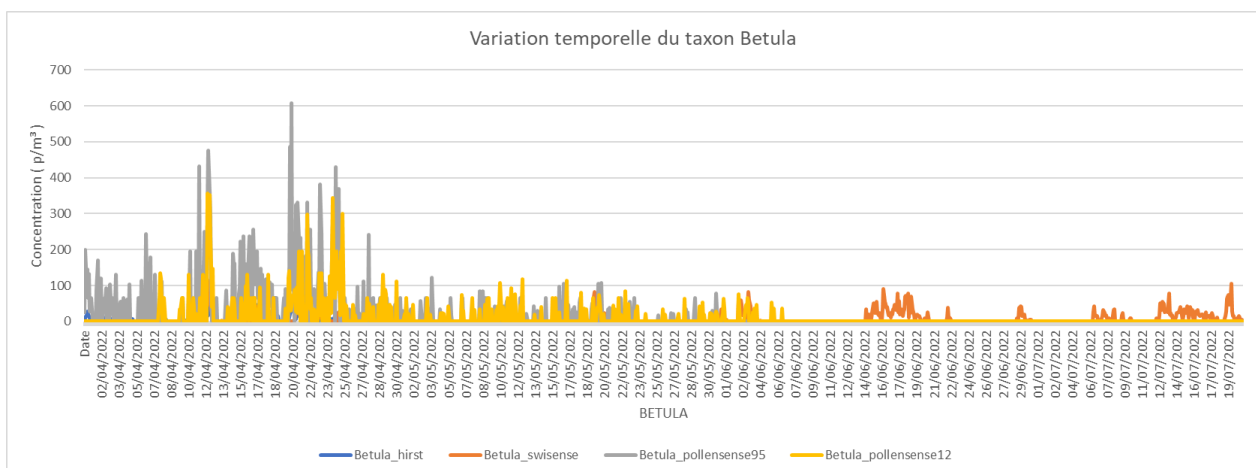


Figure 29 : Variation temporelle du Betula

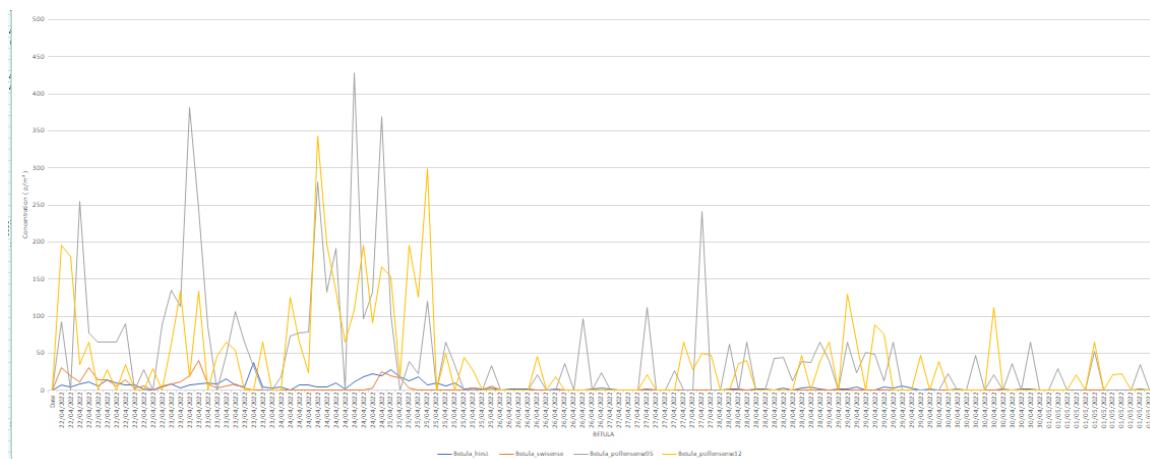


Figure 30 : Zoom sur la variation temporelle du Betula du 22 avril au 1^{er} mai

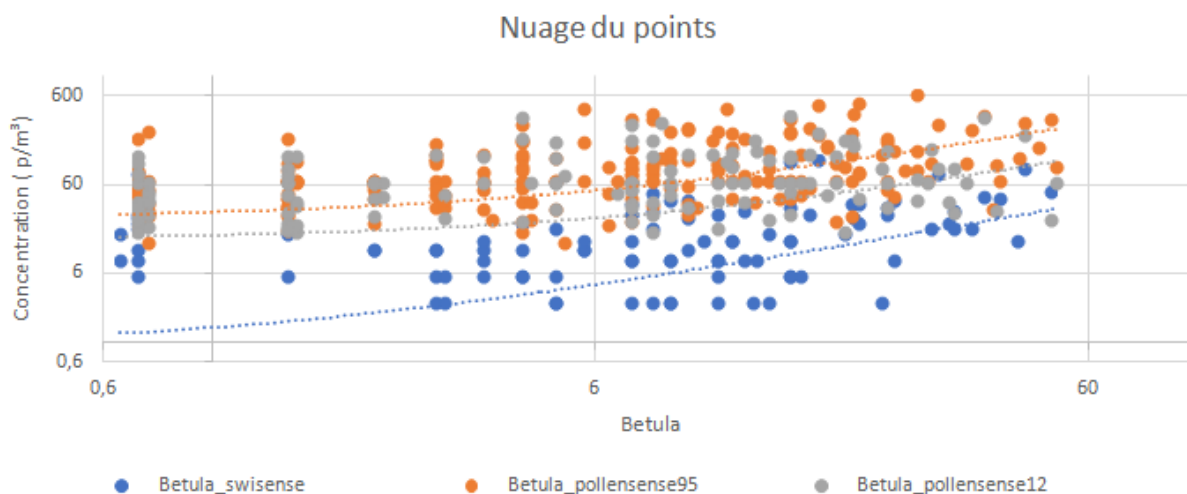


Figure 31 : Nuage de points du Betula

D'après les données du RNSA, la saison pollinique du Betula couvre la période allant de mi-mars à mi-mai. La période d'expérimentation a donc eu lieu pendant la période pollinique du Betula.

Concernant les résultats de mesure, on note une différence des intensités des concentrations, avec des concentrations pour les capteurs Hirst et Swisens beaucoup plus faibles en comparaison aux systèmes Pollensense.

Le graphique zoomé pour les journées du 22 avril au 1^{er} mai montre une variation temporelle pour l'ensemble des capteurs certainement lié à l'aspect biologique et phénologique du pollen.

Le nuage de points reflète également une surestimation des concentrations par rapport au système Hirst. On note une meilleure corrélation pour des concentrations élevées, cela est peut être lié aux facteurs de conversion appliqués qui sont éventuellement plus adaptés à des concentrations élevées en pollen.

4.1.3.2 Erreur statistique

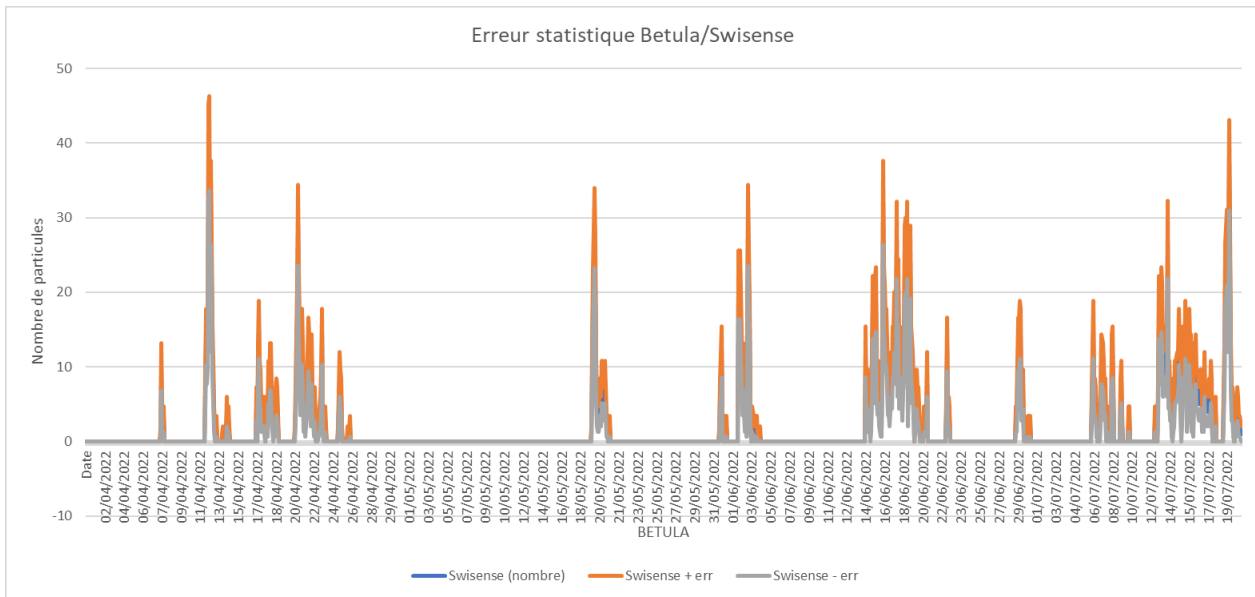


Figure 32: Erreur statistique Betula capteur SwisensePoleno

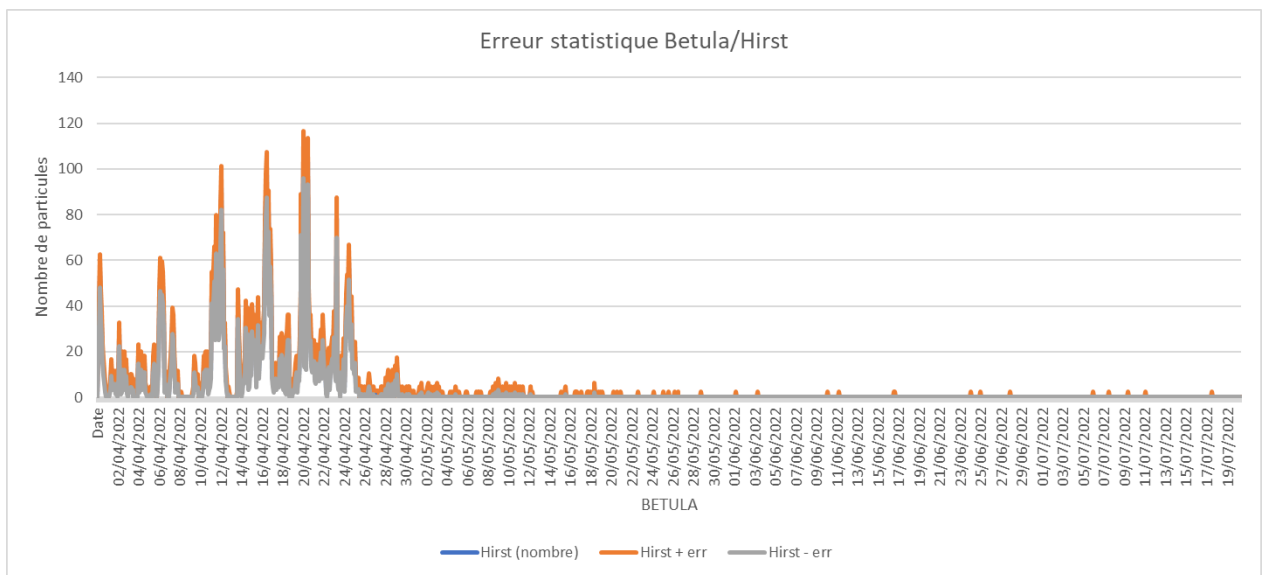


Figure 33 : Erreur statistique Betula capteur Hirst

Les graphiques représentant les erreurs statistiques montrent un bruit sur la mesure qui est négligeable.

4.1.3.3 Données statistiques

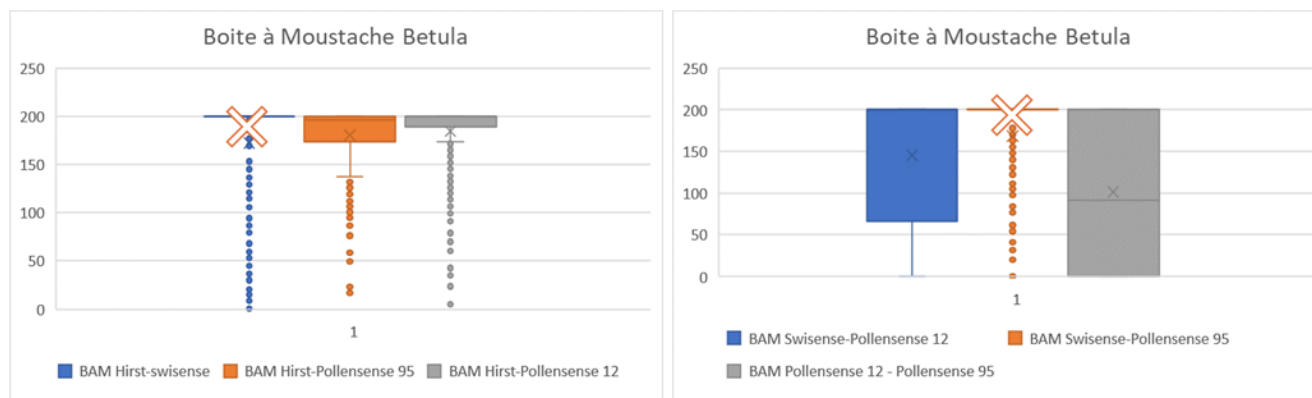


Figure 34 : Boite à Moustache Betula

Concernant les comparaisons des boites à moustaches, ces dernières sont tirées vers le haut, ce qui reflète un nombre de valeurs nulles assez important. Les résultats ne permettent donc pas de détacher des conclusions.

D'après la matrice de corrélation, le couple de capteurs le mieux corrélé est le couple Swisens/ Polensense12 avec un facteur de corrélation de 0.36.

		Matrice de Corrélation Betula			
		Hirst	Swisense	Pollensense 95	Pollensense 12
Hirst		1,00	0,28	0,40	0,19
Swisense		0,28	1,00	0,29	0,36
Pollensense 95		0,40	0,29	1,00	0,31
Pollensense 12		0,19	0,36	0,31	1,00

Figure 35 : Matrice de corrélation Betula

4.1.4 Taxon Alnus

4.1.4.1 Variation temporelle

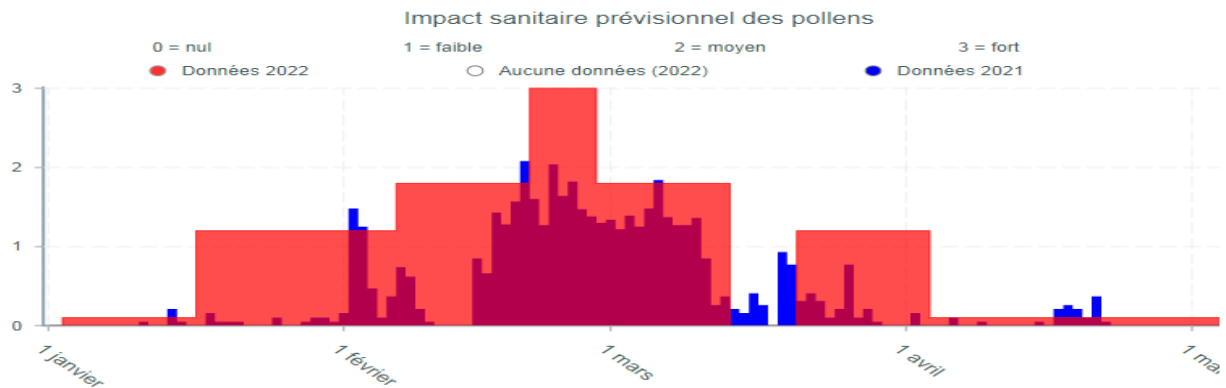


Figure 36 : Données de l'Alnus du RNSA

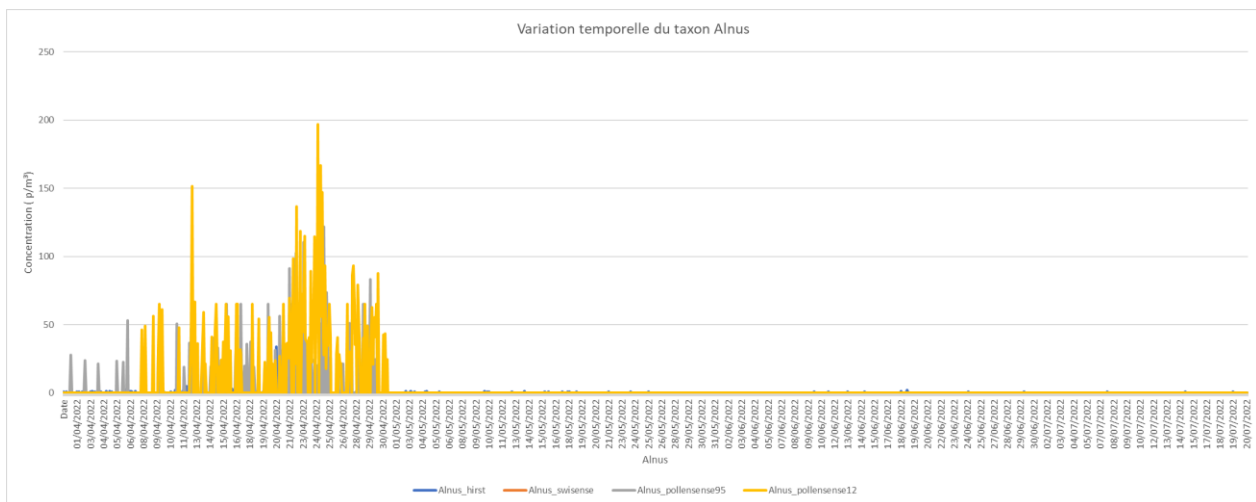


Figure 37 : Variation temporelle de l'Alnus

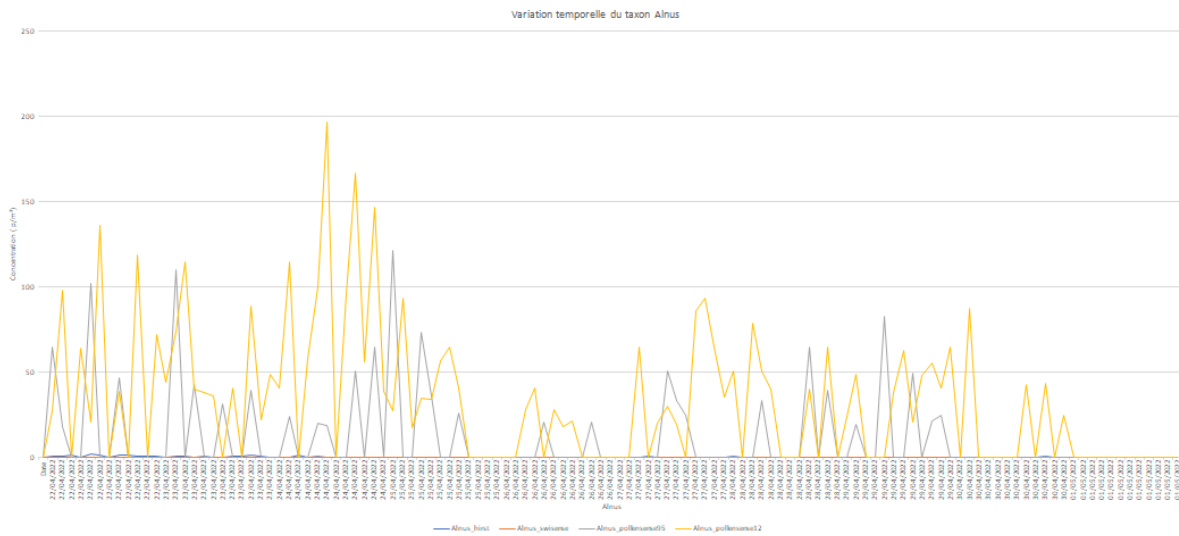


Figure 38 : Zoom sur la variation temporelle de l'Alnus du 22 avril au 1^{er} mai

D'après les données du RNSA, la saison pollinique de l'Alnus couvre la période allant de mi-janvier à fin mars. Nous sommes donc hors période pollinique pour l'Alnus pendant la période de l'expérimentation.

De manière générale on note des concentrations nulles pour le système Swisens et très peu de valeurs non nulles pour les systèmes Hirst. Seuls les capteurs Pollensense enregistrent des concentrations.

Le nuage de points n'est pas applicable pour ce cas au vu des concentrations quasi nulles du système Hirst et du faible nombre de valeurs non nulles pour l'ensemble des 3 autres capteurs.

4.1.4.2 Erreur statistique

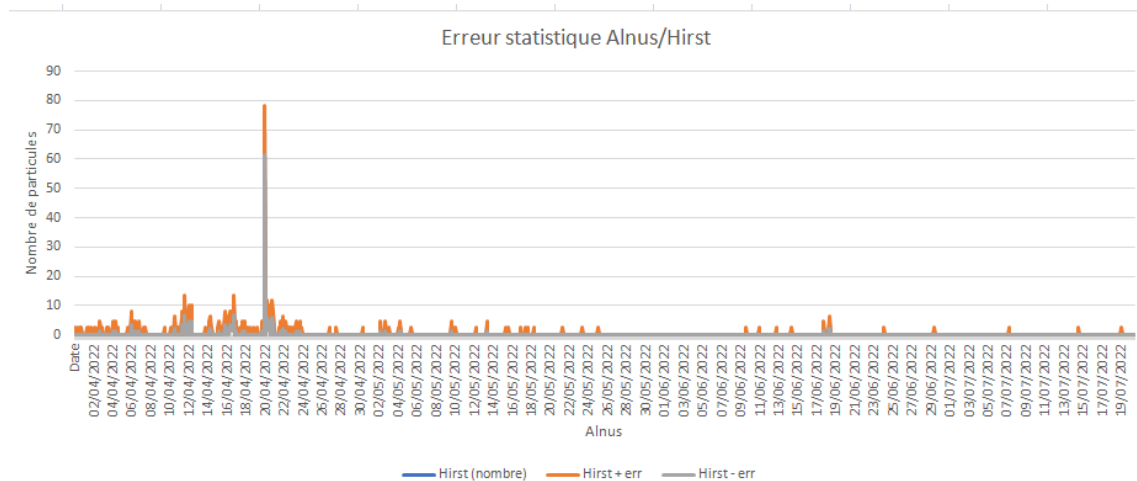


Figure 39: Erreur statistique Alnus capteur Hirst

Il n’y a pas de graphique d’erreur statistique pour le capteur Swisens au vu des valeurs nulles sur ce capteur.

Le graphique représentant les erreurs statistiques du capteur Hirst montre un bruit sur la mesure qui est négligeable.

4.1.4.3 Données statistiques

Au vu du nombre de valeurs nulles les boîtes à moustache ne sont pas applicables. Concernant la matrice de corrélation, les facteurs de corrélations sont nuls voir même négatifs pour les mêmes raisons que précédemment à savoir que la majorité des données sont nulles.

Matrice de Corrélation Alnus				
	Hirst	Swisense	Pollensense 95	Pollensense 12
Hirst	1,00	0,00	-0,17	-0,16
Swisense	0,00	1,00	0,00	0,00
Pollensense 95	-0,17	0,00	1,00	-0,12
Pollensense 12	-0,16	0,00	-0,12	1,00

Figure 40 : Matrice de corrélation de l’Alnus

4.1.5 Taxon Corylus

4.1.5.1 Variation temporelle

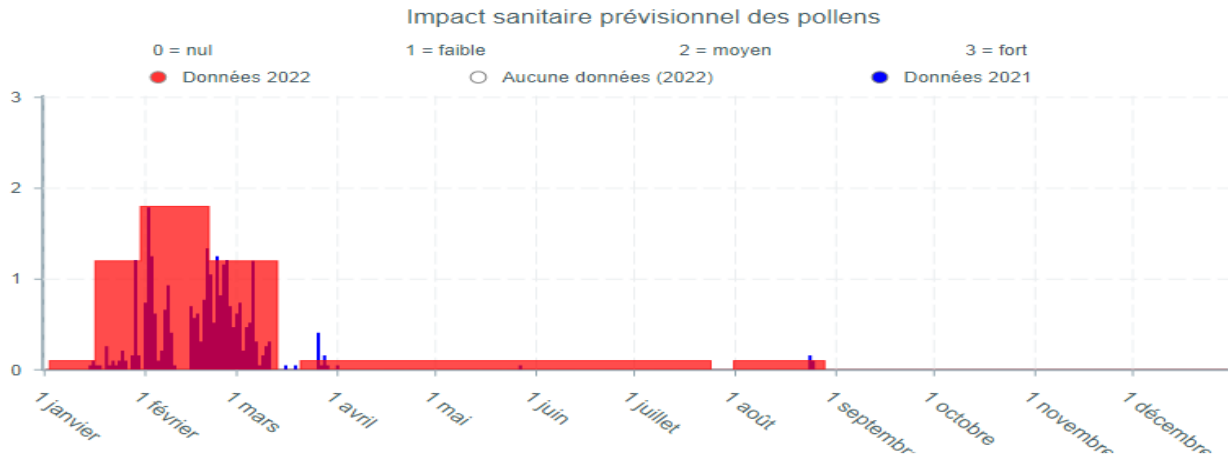


Figure 41 : Données Corylus du RNSA

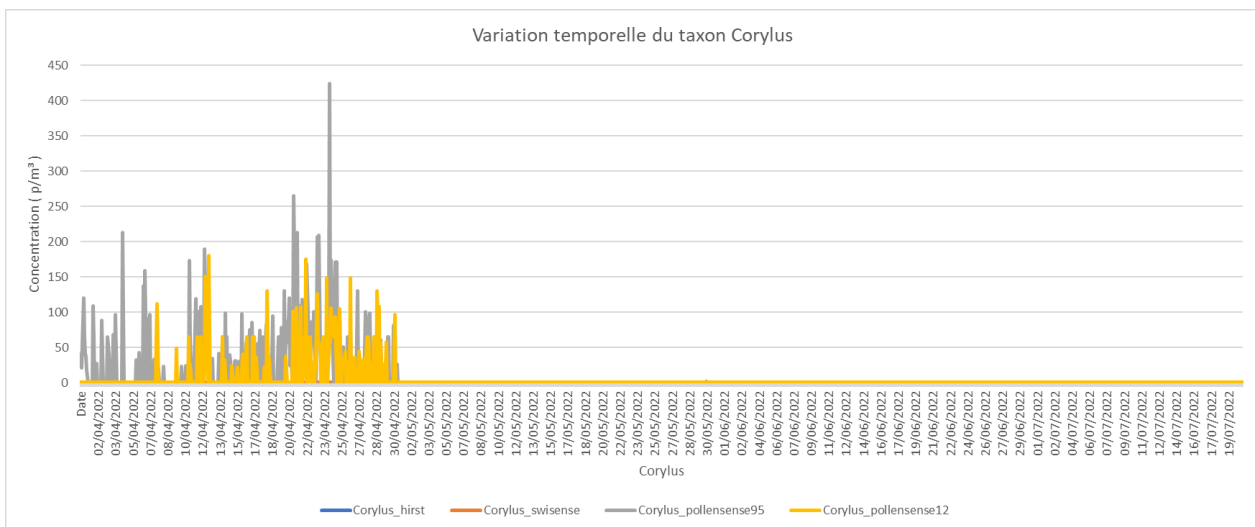


Figure 42 : Variation temporelle du Corylus

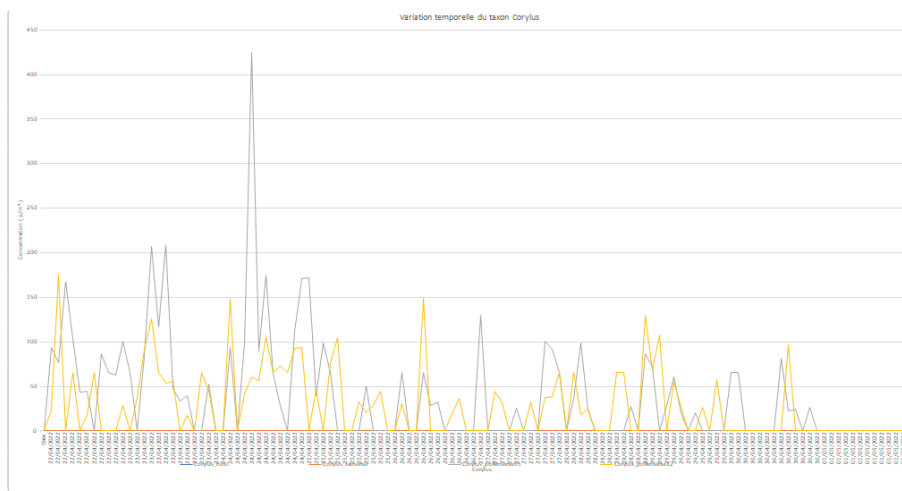


Figure 43 : Zoom sur la variation temporelle du Corylus du 22 avril au 1^{er} mai

Les données du RNSA notent une période pollinique du Corylus se terminant mi- mars. Nous sommes donc hors période pollinique pour le Corylus pendant la période de l'expérimentation.

Concernant les résultats, on note des concentrations nulles pour le système Swisens et très peu de valeurs non nulles pour les systèmes Hirst. Seuls les capteurs Pollensense enregistrent des concentrations.

Le nuage de points n'est pas applicable pour ce cas au vu des concentrations quasi nulles du système Hirst et du faible nombre de valeurs non nulles pour l'ensemble des 3 autres capteurs

4.1.5.2 Erreur statistique

Le calcul de l'erreur statistique n'est pas applicable pour le capteur Swisens au vu des concentrations nulles sur l'ensemble de la période d'expérimentation.

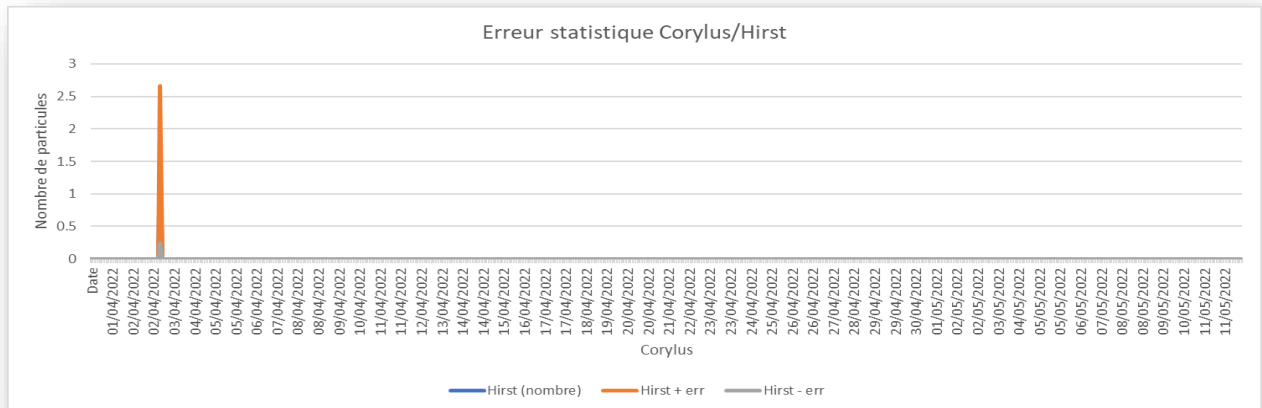


Figure 44 : Erreur statistique Corylus capteur Hirst

Le graphique représentant les erreurs statistiques sur le capteur Hirst montre un bruit sur la mesure qui est négligeable.

4.1.5.3 Données statistiques

Au vu du nombre de valeurs nulles les boîtes à moustache ne sont pas applicables. Concernant la matrice de corrélation, les facteurs de corrélations sont nuls voir même négatifs pour les mêmes raisons que précédemment à savoir que la majorité des données sont nulles.

Matrice de Corrélation Corylus				
	Hirst	Swisense	Pollensense 95	Pollensense 12
Hirst	1,00	0,00	0,00	0,00
Swisense	0,00	1,00	0,00	0,00
Pollensense 95	0,00	0,00	1,00	0,13
Pollensense 12	0,00	0,00	0,13	1,00

Figure 45 : Matrice de corrélation Corylus

4.1.6 Taxon Platanus

4.1.6.1 Variation temporelle

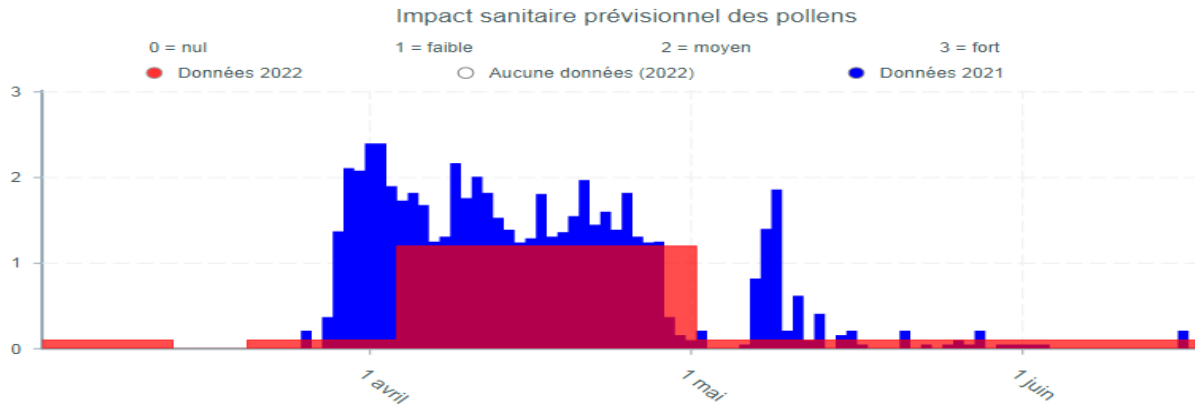


Figure 46 : Données Platanus du RNSA

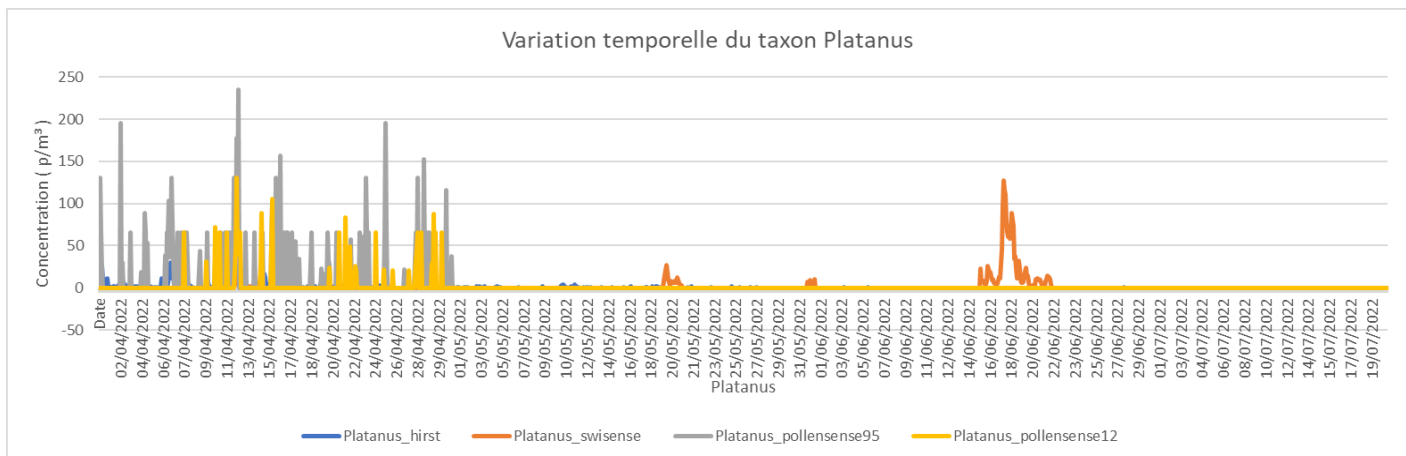
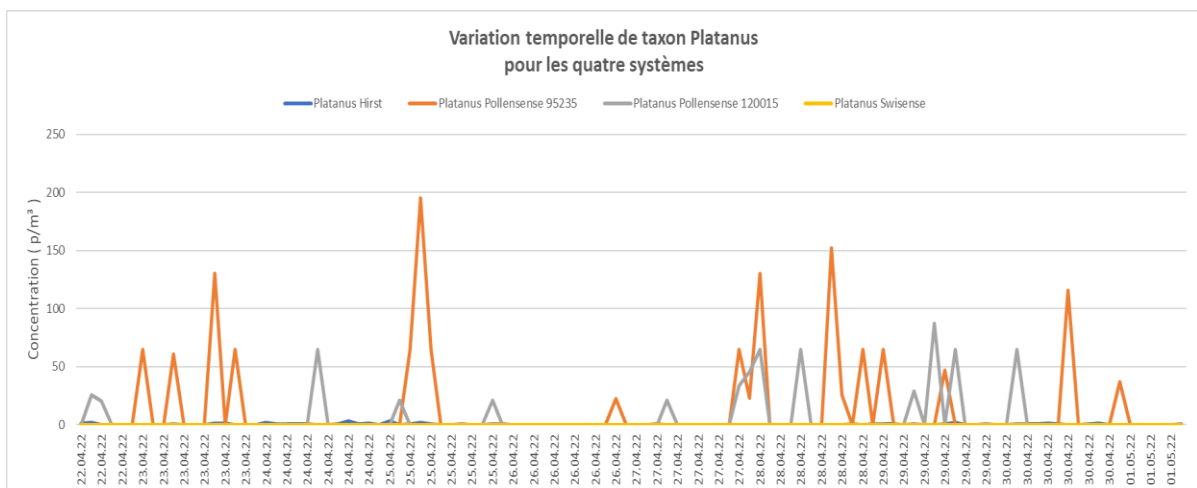


Figure 47 : Variation temporelle du Platanus



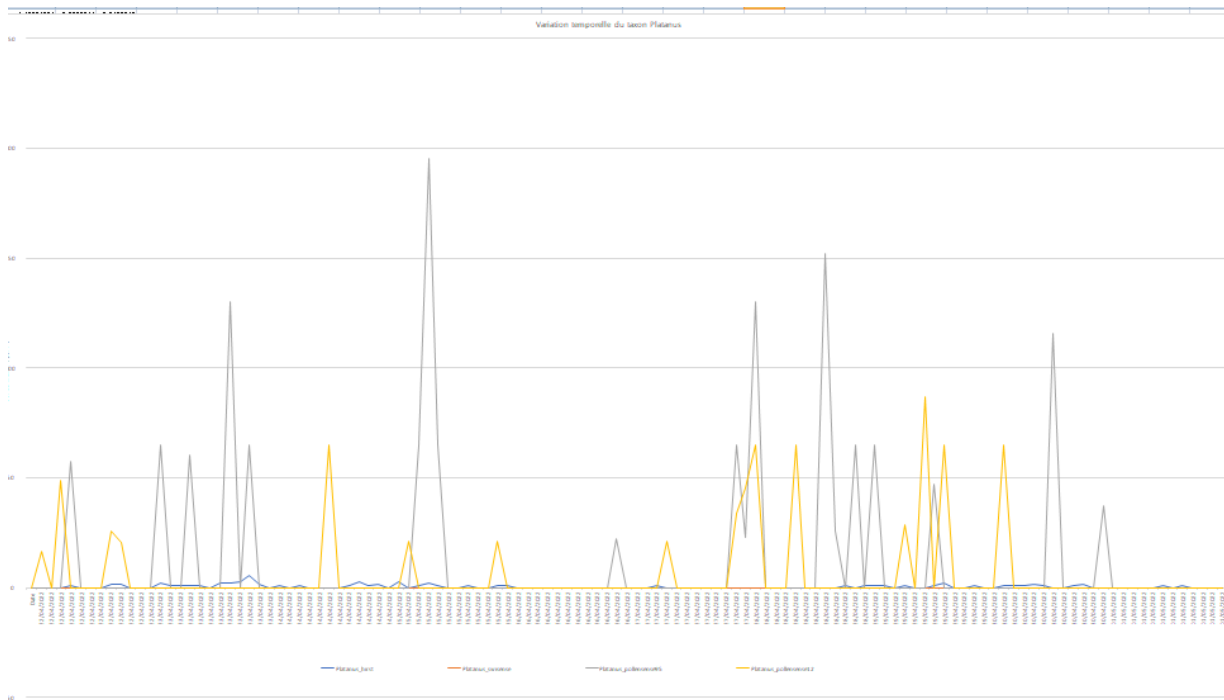


Figure 48 : Zoom sur la variation temporelle du Platanus du 22 avril au 1^{er} mai

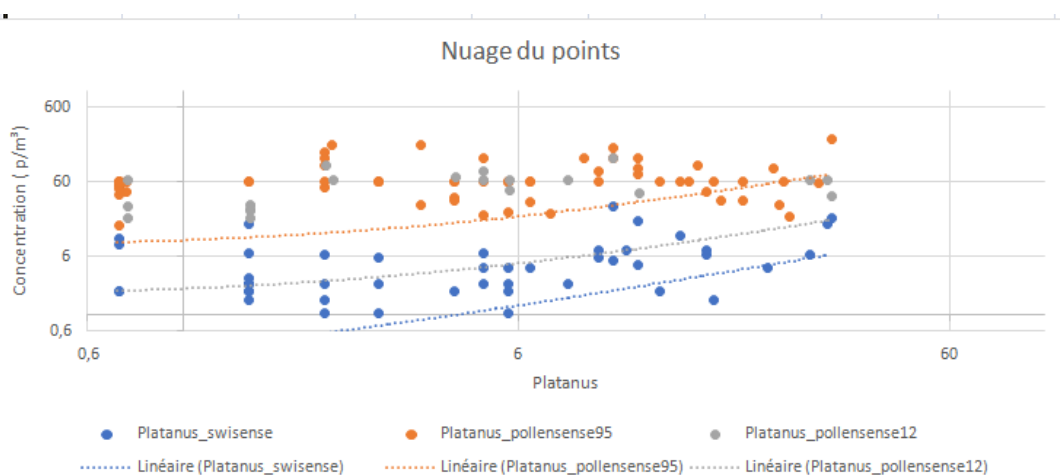


Figure 49 : Nuage de points du Platanus

Les données du RNSA notent une période pollinique couvrant le mois d'avril. Les résultats de mesure sur l'ensemble des 3 capteurs Hirst, Pollensense12 et Pollensense95 concordent avec cette saison pollinique

D'après les résultats de mesure, on note très peu de valeurs non nulles pour le capteur Swisens et des valeurs faibles pour le capteur Hirst en comparaison avec ceux de Pollensense

Le nuage de points montre également une absence de regroupement des points au vu d'une majorité de valeurs nulles.

4.1.6.2 Erreur statistique

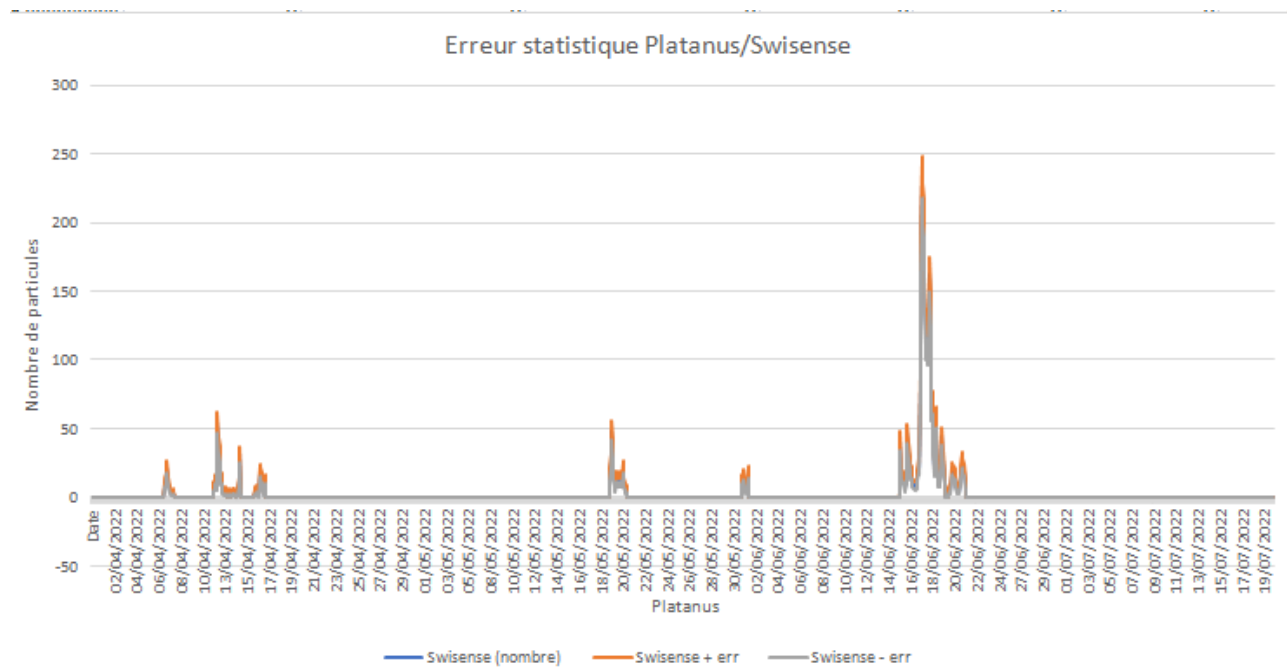


Figure 50 : Erreur statistique Platanus capteur SwisensPoleno

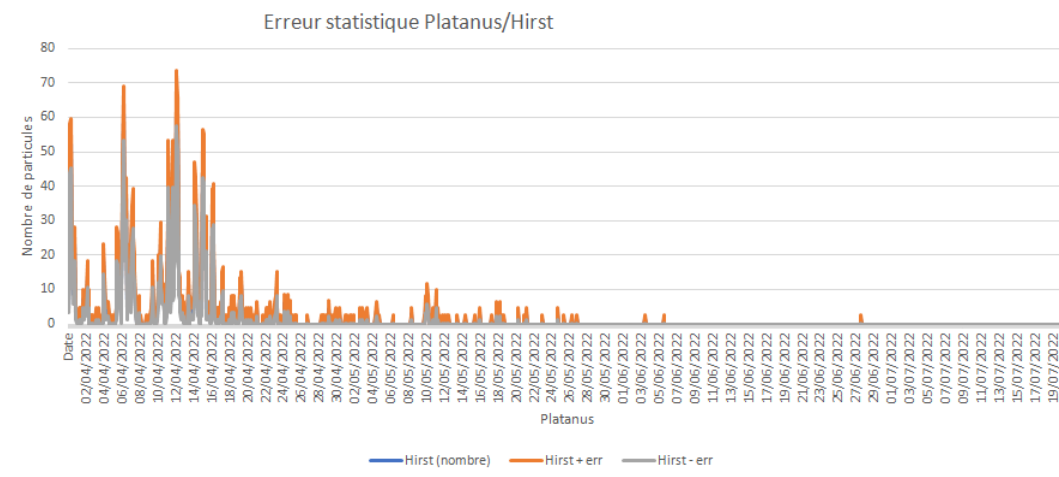


Figure 51 : Erreur statistique Platanus capteur Hirst

Les graphiques représentant les erreurs statistiques montrent un bruit sur la mesure qui est négligeable. La variabilité observée est liée à la variabilité naturelle de l'espèce.

4.1.6.3 Données statistiques

Au vu du nombre de valeurs nulles, les boîtes à moustache ne sont pas applicables. Concernant la matrice de corrélation, les capteurs les mieux corrélés sont Swisens et pollensense12 avec un facteur de 0.33.

Matrice de Corrélation Platanus				
	Hirst	Swisense	Pollensense 95	Pollensense 12
Hirst	1,00	0,29	0,15	0,05
Swisense	0,29	1,00	0,14	0,33
Pollensense 95	0,15	0,14	1,00	0,02
Pollensense 12	0,05	0,33	0,02	1,00

Figure 52 : Matrice de corrélation Platanus

4.2 ANALYSE POUR LA PERIODE DU 12 MAI AU 15 SEPTEMBRE DES 3 CAPTEURS : POLLENSENSE 12/95 ET SWISENS

Ce chapitre présente les graphiques obtenus pour l'ensemble des capteurs pour la période du 4 août à septembre.

4.2.1 Taxon Quercus

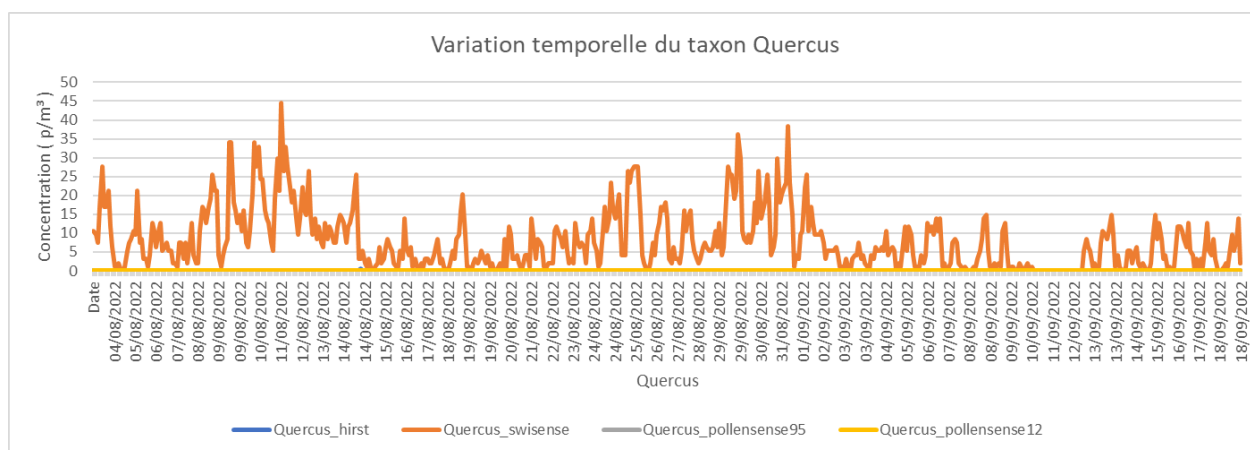


Figure 53 : Variation temporelle du Quercus aout- septembre

Seul le capteur swisens enregistre des concentrations sur l'ensemble de la période.

4.2.2 Taxon Fraxinus

Les concentrations sont nulles pour l'ensemble des capteurs.

4.2.3 Taxon Betula

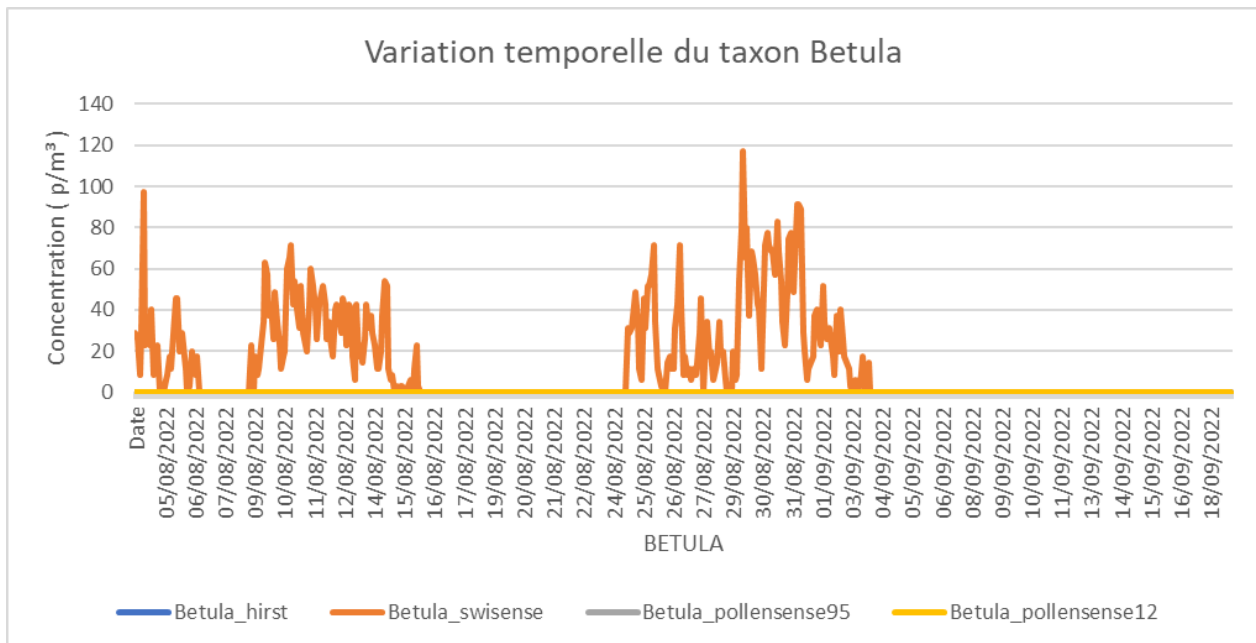


Figure 54 : Variation temporelle du Betula aout- septembre

Seul le capteur swisens enregistre des concentrations sur l'ensemble de la période.

4.2.4 Taxon Alnus

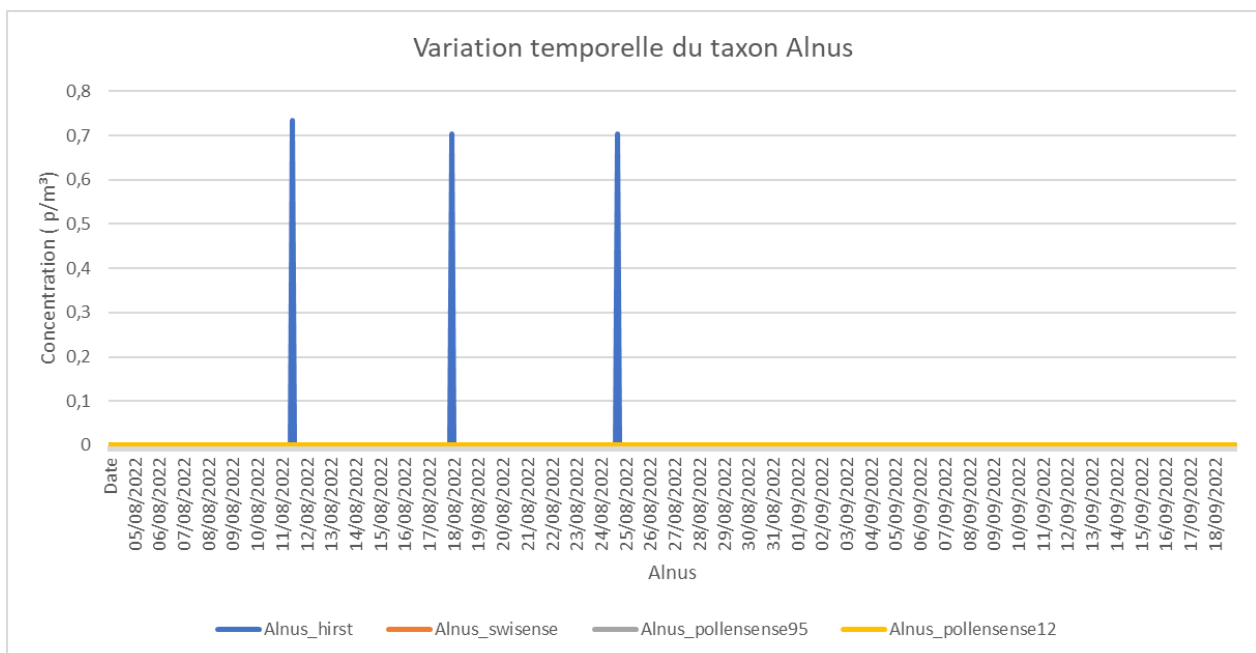


Figure 55 : Variation temporelle Alnus aout- septembre

Seul le capteur Hirst enregistre quelques pics.

4.2.5 Taxon Corylus

Les concentrations sont nulles pour l'ensemble des capteurs.

4.2.6 Taxon Platanus

Les concentrations sont nulles pour l'ensemble des capteurs.

5 ANALYSE DES COMPOSANTES PRINCIPALES

5.1 INTRODUCTION

La prévision des concentrations en pollens repose sur différents facteurs. Outre les aspects météorologiques, les aspects phénologiques (influence climatique sur le développement saisonnier des plantes) et biologiques des plantes sont également un aspect non négligeable quant à la prévision des concentrations en pollen.

Les travaux préliminaires menés dans ce cadre ont consisté à évaluer un lien entre les paramètres qui ont pu être mesurés pendant cette expérimentation qui sont les concentrations en pollen et les conditions météorologiques. Cette comparaison entre les différents paramètres repose sur une méthodologie d'analyse de composantes principales (ACP).

Pour une évaluation plus complète, ces comparaisons devront être couplées avec l'aspect biologique de chaque espèce.

5.2 OUTIL D'ANALYSE

L'analyse des composantes principales peut se faire avec l'outil RStudio qui est un langage de programmation utilisé pour le traitement de données et de l'analyse statistique. Il s'agit d'un environnement de développement gratuit, multiplateforme pour R et disponible sous licence libre. Il est disponible sous le lien suivant : <https://www.rstudio.com/products/rstudio/#rstudio-desktop>

Cet outil possède une puissance de calcul statistique avancée et une analyse multifactorielle sur la base d'une intelligence artificielle. Il comprend des packages d'analyses permettant :

- l'analyse exploratoire multidimensionnelle de données via **Factominer**,
- l'amélioration facile et interactive des graphiques pour les rendre beaucoup plus lisibles via **Factoshiny**.

Il permet également de générer automatiquement un rapport statistique

5.3 STRATEGIE D'ANALYSE

Une première analyse a été réalisée sur la base des données disponibles au moment de la réalisation du projet.

Pour cela il a été décidé de se baser sur :

- Les données du Système Hirst qui est le système jugé le plus aboutit au moment de l'analyse ACP
- Les concentrations du taxon Quercus qui est l'espèce donnant le plus de résultats non nuls pendant la période de mesure de l'expérimentation.
- Une période temporelle allant de mi-avril à mi-mai correspondant à la période pollinique du Quercus selon le RNSA.

Les données météorologiques Météo France du site ne sont pas disponibles. Des données d'extraction ont donc été récupérées sur un point de grille le plus proche du lieu d'expérimentation (grille de 12 km).

L'analyse météorologique se base sur cinq variables météorologiques à savoir :

- La composante horizontale de la vitesse U
- La composante verticale de la vitesse V
- La température
- La précipitation
- Le rayonnement solaire

5.4 RESULTATS

L'outil permet d'extraire des cercles de corrélation sur les deux dimensions jugées les plus pertinentes après avoir réduit le nombre de variables.
Sur le jeu de données utilisés pour cette analyse, le cercle de corrélation obtenu est disponible ci-dessous.

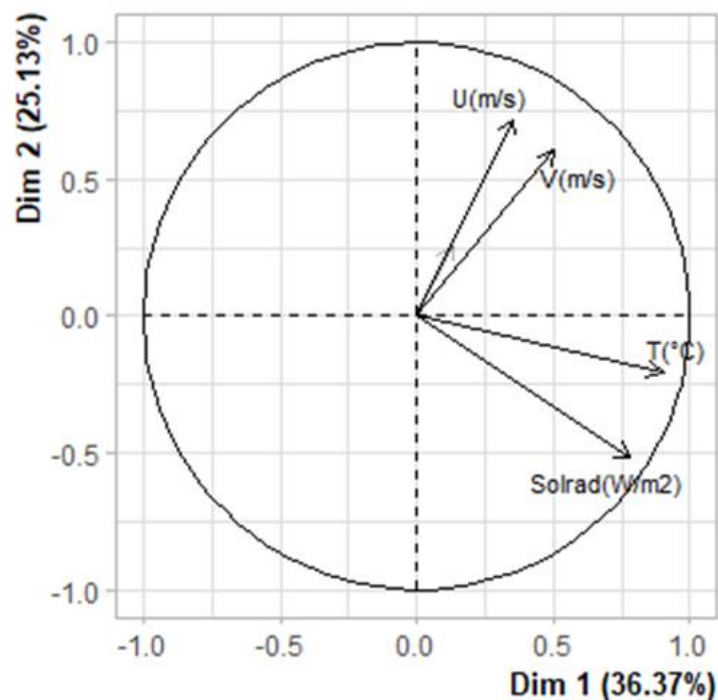


Figure 56 : Cercle de corrélation

Ce qu'on peut déduire de ce cercle de corrélation est que :

- La vitesse est corrélée avec l'axe vertical
- La température et le rayonnement sont corrélés avec l'axe horizontal
- Les deux variables météorologiques les plus optimales sont la composante U de la vitesse et la température

Le détail de cette analyse est disponible dans le document envoyé au client nommé « **Pollen-Rapport_ACP.ppt** »

Les données disponibles ne permettent pas de mener l'analyse ACP dans sa globalité et tirer des conclusions générales.

Au regard du retour d'expérience sur la campagne de mesure, l'analyse ACP devrait :

- Utiliser sur une plus grande série de données sur plusieurs saisons afin de vérifier la reproductibilité des phénomènes liée aux paramètres météorologiques.
- Utiliser des données météorologiques locales.
- Corréler et/ou coupler avec d'autres bases d'informations concernant l'aspect biologique et phénologique du taxon étudié.

6 CONCLUSION ET PERSPECTIVES

Pour répondre aux besoins de l'analyse des systèmes de mesure, Fluidyn a réalisé des analyses comparatives détaillées des concentrations en taxons issues de chacun des 4 systèmes de mesure positionnées au niveau de la gare Saint-Lazare et qui sont :

- Système SwisensPoleno Mars
- 2 systèmes Pollensense nommés 95235 et 120015
- Système Hirst considéré comme le système de référence

Cette analyse porte sur les taxons communs à l'ensemble des systèmes soit un ensemble de six taxons. Certains pollens comme l'ambrosie, pollen considéré comme hautement allergisante et faisant l'objet d'une attention particulière, ne figure pas dans la liste des espèces mesurées par Swisens et Pollensense.

L'analyse couvre la période d'expérimentation allant du 31 mars à septembre 2022. L'analyse des données a montré que la période d'expérimentation ne couvre pas l'ensemble de la période pollinique pour les taxons communs. Il serait donc intéressant de renouveler l'expérimentation sur une période plus longue afin d'affiner les résultats.

Afin d'automatiser au maximum le travail d'analyse comparative, un outil a été développé afin de réaliser l'ensemble des calculs nécessaires à savoir

- Les graphiques des séries temporelles
- Les nuages de points
- Les erreurs statistiques
- Les boîtes à moustache
- Les matrices de corrélations

Cet outil a été développé et pensé sur la base des contraintes relevées lors de l'expérimentation. Ces contraintes sont liées au nombre et type de taxons, nombre de capteurs et repose sur le système de référence Hirst.

Une des évolutions possibles de cet outil pourrait impliquer l'indépendance du nombre et type de taxons et du nombre de capteurs.

Afin d'enrichir l'analyse comparative, il serait intéressant de poursuivre l'expérimentation sur une période plus étendue et de la renouveler. Cela permettrait de couvrir la période pollinique de l'ensemble des taxons et alimenter en base de données le modèle ACP pour de meilleure corrélation des données météo avec les concentrations en Pollen.