

Rapport de stage de Master 2 professionnel
Gestion de l'Environnement Industriel
Université Grenoble I – UFR de physique
2005-2006

Représentativité statistique du réseau national de surveillance aérobiologique (RNSA) en Rhône-Alpes. Faisabilité d'une modélisation.

Structure d'accueil : ASCOPARG
Tuteur : Nicolas Vigier



Camille Rieux
camille_rieux@yahoo.fr

SOMMAIRE

I. OBJECTIF DE L'ETUDE.....	3
II. GENERALITES: AEROBIOLOGIE – ALLERGOLOGIE.....	4
1. Le pollen.....	4
2. Les facteurs météorologiques qui influencent le contenu pollinique de l'air	5
3. Les pollens allergisants	5
4. Mécanismes et symptômes de l'allergie au pollen.....	5
5. Prévention - traitement	5
6. Métrologie	6
6.1. objectifs des observatoires sanitaires	6
6.2. identification des grains de pollen.....	6
6.3. les capteurs	6
7. Réglementation sur les pollens.....	8
III. ETUDE DU PLAN D'ECHANTILLONNAGE DU RESEAU	8
1. Stratégie générale	8
2. Cadre de l'étude statistique.....	9
2.1. base de donnée.....	9
2.2. prétraitement des données	9
2.3. approche statistique	10
2.4. définition des limites de l'étude	11
3. Résultat – Discussion	12
3.1. étude 1 : année 2005	12
3.2. étude 2 : période 2003-2005	19
4. Conclusions – possibilités d'optimisations.....	20
IV. MODELISATION STATISTIQUES.....	22
1. Stratégie générale de l'exercice de modélisation.....	22
2. Cadre de l'étude et méthode.....	23
2.1. définition des limites de l'étude	23
2.2. modélisation par interpolation.....	23
2.3. modélisation par ajustement de la phase et/ou de l'amplitude.....	24
3. Résultats – Discussion.....	26
3.1. interpolation.....	26
3.2. modèles basés sur l'ajustement de la phase et/ou de l'amplitude.....	28
3.3. Mise en œuvre des différents modèles	31
4. Conclusions	34
V. CONCLUSION GENERALE	35

ABREVIATIONS

Taxon (au pluriel taxa) : les différentes "catégories" de pollen sont désignées par le terme taxon. Le nom d'un taxon désigne l'espèce, le genre ou la famille de la plante qui produit le pollen selon le niveau systématique d'identification

DPP : degré de proximité pollinique entre deux sites pour un taxon donné

DPP_{mt} : degré de proximité pollinique entre deux sites moyenné pour tous les taxons considérés dans l'étude. "mt" signifiant "moyenne taxon".

DPP-MD₂₀₀₅ : degré de proximité pollinique entre les décours temporels modélisés (M, modèle) et ceux qui ont été effectivement enregistrés (D, donnée) pour l'année 2005.

DPP-MD₂₀₀₃₋₂₀₀₄ : degré de proximité pollinique entre les décours temporels modélisés (M, modèle) et ceux qui ont été effectivement enregistrés (D, donnée). Les valeurs obtenues pour l'année 2003 et 2004 ont été moyennées.

Int : modèle basé sur l'interpolation de la réponse des sites voisins

IntA : modèle basé sur l'ajustement de l'amplitude sur la base de l'enregistrement modélisé par interpolation

SiteA : modèle basé sur l'ajustement de l'amplitude sur la base de l'enregistrement d'un site donnée du réseau

IntAP : modèle basé sur l'ajustement de l'amplitude et de la phase sur la base de l'enregistrement modélisé par interpolation

SiteAP : modèle basé sur l'ajustement de l'amplitude et de la phase sur la base de l'enregistrement d'un site donné du réseau

I. OBJECTIF DE L'ETUDE

Les pollinoses : un problème naissant de santé publique

La prévalence des pathologies allergiques et des pollinoses¹ en particulier n'a cessé d'augmenter au cours des dernières décennies dans les pays industrialisés. Actuellement, 15% de la population française présenteraient une allergie aux pollens. Cette affection se traduit généralement par une rhinite allergique, souvent associée à une conjonctivite et parfois à un asthme. La pollinose peut donc s'avérer très invalidante pour les personnes sensibilisées, en particulier si elle est à l'origine de troubles du sommeil et de céphalée. Outre la souffrance et le handicap des personnes allergiques, les pollinoses sont à l'origine d'un préjudice financier élevé pour la collectivité lié aux coûts de santé directs (consultations, traitements et examens) et indirects (absentéisme professionnel). Les problèmes d'allergie liés à la présence de pollen dans l'air, constituent donc une préoccupation croissante en matière de santé publique.

La surveillance du contenu pollinique de l'air

Bien qu'il n'y ait pas d'obligation réglementaire relative à la surveillance des pollens, différents réseaux météorologiques assurent la mesure du contenu pollinique de l'air. En France deux méthodologies coexistent, l'une est fondée sur le principe de la filtration (méthode de Cour) et l'autre sur le principe de l'aspiration (méthode de Hirst). Chacune de ces méthodes présente des avantages spécifiques en particulier en ce qui concerne l'incertitude de mesure et la finesse du pas de mesure. Il existe trois réseaux basés sur des capteurs de type Cour localisés sur le pourtour méditerranéen (3 stations), en outre-mer (5 stations) et dans la vallée du Rhône (5 stations). Le Réseau National de Surveillance Aérobiologique (RNSA, association loi 1901) compte quant à lui 54 capteurs de Hirst répartis sur l'ensemble du territoire métropolitain. L'étude présentée dans ce rapport ne concerne que les données collectées par les stations du RNSA de la région Rhône-Alpes et de ses environs.

Outre son activité purement météorologique, le RNSA assure un rôle d'information et diffuse chaque semaine un bulletin allergeo-pollinique national (www.rnsa.fr). Les informations relatives au déroulement de la pollinisation des différentes espèces permettent aux personnes allergiques d'adapter leur comportement et leurs traitements médicamenteux aux conditions du moment. De plus, les informations collectées chaque année par le RNSA constituent une base de données qui peut être exploitée dans le cadre d'études spécifiques portant par exemple sur la modification de l'aire de répartition d'une espèce comme l'ambrosie ou la caractérisation de la relation entre les facteurs météorologiques et les concentrations polliniques.

L'ASCOPARG (ASsociation pour le CONtrôle et la Préservation de l'Air en Région Grenobloise) est l'une des 40 AASQA (Associations Agréées de Surveillance de la Qualité de l'Air) qui forment le réseau ATMO. Les missions réglementaires des AASQA sont la surveillance des polluants réglementés (ozone, dioxyde d'azote, dioxyde de soufre et poussières, etc.) et la diffusion d'informations relatives à la qualité de l'air. Néanmoins, l'activité de l'ASCOPARG dépasse largement ses obligations réglementaires et l'association s'intéresse notamment aux pollens. En effet l'ASCOPARG réalise, pour le compte du RNSA, l'analyse du contenu pollinique de l'air au niveau de deux sites de mesure du département de l'Isère (Grenoble et Roussillon), et chaque semaine le bulletin allergeo-pollinique régional élaboré par le RNSA est diffusé sur le site internet des AASQA de la région (www.atmo-rhonealpes.org). C'est donc au titre d'acteur du RNSA que l'ASCOPARG souhaitait effectuer l'étude présentée dans ce rapport.

¹ Le terme pollinose désigne l'ensemble des manifestations cliniques induites par l'allergie aux pollens.

Objectifs de l'étude

Le réseau métrologique du RNSA s'est étendu progressivement au cours des vingt dernières années. La stratégie d'échantillonnage qui a prévalu jusqu'à présent consiste à installer des stations de mesure dans les principales agglomérations. Afin de valoriser les efforts et l'argent investis dans le développement et la mise en œuvre au quotidien du réseau métrologique du RNSA, une étude portant sur le réseau métrologique de la région Rhône-Alpes et de ses environs a été réalisée sur proposition de l'ASCOPARG.

Cette étude comporte deux objectifs:

Objectif 1. évaluer la représentativité du réseau métrologique, c'est à dire identifier les éventuelles redondances ou carences du plan d'échantillonnage, et le cas échéant faire des propositions visant à l'optimiser (déplacement, ajout ou élimination de sites de mesure).

Objectif 2. tester la possibilité de modéliser les données fournies par les capteurs des différents sites dans la perspective de développer une stratégie couplant métrologie de constat et modélisation.

L'étude présentée a été réalisée à l'aide d'un programme développé sous matlab pour les besoins de cette étude. Cet outil informatique est piloté par l'intermédiaire de boîtes de dialogue, son utilisation ne nécessite donc pas de connaissance en programmation. Ainsi, des études similaires, visant à mesurer la validité de la localisation de nouveaux sites ou à réactualiser les paramètres des modèles, pourront être réalisées à l'ASCOPARG ultérieurement. Ce document tient donc lieu à la fois de rapport d'étude et de démonstration de l'usage qui peut être fait du programme.

II. GENERALITES: AEROBIOLOGIE – ALLERGOLOGIE

1. Le pollen

La morphologie du grain de pollen est caractéristique de l'espèce de plante qui le produit, sa taille varie de 5 à 250 μm . Le grain de pollen est l'élément reproducteur produit par les organes mâles des plantes supérieures (anthères des étamines). La déhiscence de l'anthère conduit à la libération des grains de pollens qui doivent parvenir à l'organe reproducteur femelle (style) pour que se déroule la fécondation. Selon les espèces, le transport des grains de pollens est assuré par les insectes (plantes entomophiles) ou par le vent (plantes anémophiles). D'une manière générale, les plantes entomophiles ont des fleurs colorées, parfumées, riches en nectar et produisent peu de pollen. A l'inverse, les plantes anémophiles produisent une grande quantité de grains de pollens (un seul pied d'ambrosie peut produire jusqu'à $2,5 \cdot 10^9$ grains en une journée), ils sont de petite taille (généralement de 20 à 60 μm) et sont donc susceptibles d'être transportés sur de grandes distances (les grains de noisetier parcourent couramment 250 km). Ces propriétés assurent le succès de la reproduction des plantes anémophiles en dépit de la faible probabilité de rencontre du grain de pollen avec le style.

Chaque espèce végétale pollinise au cours d'une période spécifique. Par exemple, les arbres tels que le cyprès et le bouleau émettent des grains de pollen dès la fin de l'hiver et au cours du printemps, les graminées pollinisent de mai à juillet et l'ambrosie, espèce tardive, libère son pollen au cours du mois de septembre. Il convient toutefois de noter que la période de pollinisation, de même que la quantité de grain produite, bien qu'étant des caractéristiques intrinsèques d'une espèce, sont conditionnées par les conditions météorologiques rencontrées au cours de la période de développement floral et au moment de la libération des grains.

2. Les facteurs météorologiques qui influencent le contenu pollinique de l'air

La production, la libération et le transport des grains de pollen déterminent la concentration en grain de pollen à un endroit et à un instant donné. Ces trois paramètres sont influencés par des facteurs météorologiques. Schématiquement, la production de pollen est influencée par les conditions de température et d'ensoleillement rencontrées au cours de la période de développement floral. La libération des grains dans l'atmosphère est conditionnée par la pression atmosphérique, l'humidité et la température. Le transport des grains de pollen est déterminé par les mouvements d'air horizontaux et verticaux. De plus, la pluie précipite les grains de pollen au sol et abat efficacement les concentrations polliniques. La variabilité interannuelle des conditions météorologiques entraîne donc une variabilité considérable des paramètres qui caractérisent la saison pollinique. Ainsi, d'une année à l'autre, la date de début de pollinisation d'une espèce donnée peut varier de plus de 25 jours et la production pollinique annuelle fluctuer dans un rapport de 1 à 10. Il est donc essentiel, afin d'aider les patients atteints de pollinose, que la saison pollinique soit caractérisée chaque année.

3. Les pollens allergisants

L'allergénicité des grains de pollen varie d'une espèce à l'autre. Elle est principalement déterminée par le nombre et la nature des allergènes (molécules allergisantes) que présente le grain de pollen ainsi que par la concentration de grains de pollen présents dans l'atmosphère. Ainsi, les espèces anémophiles représentent la majorité des espèces responsables de pollinoses puisqu'elles libèrent de grandes quantités de grains de pollen, transportée sur de grandes distances et qui présentent donc des concentrations élevées dans l'atmosphère, y compris en milieu urbain. Des arbres tels que le bouleau et le cyprès, et des herbacées telles que les graminées et l'ambrosie comptent parmi les plantes à l'origine des pollens les plus allergisants.

4. Mécanismes et symptômes de l'allergie au pollen

La libération des allergènes qui intervient lorsque le grain de pollen se dépose sur les yeux ou sur les muqueuses du tractus respiratoire produit une inflammation locale de ces régions. La pollinose se manifeste donc généralement par une rhinite² allergique dont les trois symptômes sont : les éternuements à répétition, la rhinorrhée ("nez qui coule") et l'obstruction nasale. La rhinite pollinique est très souvent associée à une conjonctivite³ et parfois à un asthme. La survenue de l'allergie au pollen peut être influencée par différents facteurs : le terrain génétique, la présence simultanée de plusieurs pollens allergisants, l'existence de réactions croisées entre les pollens de la même famille ou avec certains aliments ou encore la pollution atmosphérique.

5. Prévention - traitement

Trois approches permettent de prévenir ou de traiter la pollinose : l'éviction, les traitements médicamenteux et la désensibilisation. L'**éviction** est la méthode de prévention de choix mais est difficile à mettre en œuvre. Elle consiste à limiter autant que possible l'exposition aux grains de pollen durant la période de pollinisation des pollens déclenchant l'allergie. Les **médicaments anti-allergiques** permettent de diminuer l'intensité des symptômes et de

² inflammation de la muqueuse nasale

³ inflammation oculaire

combattre l'inflammation causée par l'allergie, ils contrôlent efficacement rhinite et conjonctivite allergiques. Néanmoins ces thérapeutiques ne traitent pas la cause de l'allergie, leurs effets ne durent donc que le temps de leur utilisation et n'ont pas de bénéfice à long terme.

Dans les cas où le patient connaît les pollens auxquels il est allergique, ces deux approches peuvent être ajustées en fonction des informations relatives à la progression de la pollinisation des pollens concernés, diffusées par les réseaux de surveillance aérobiologique. En cas d'échec ou d'insuffisance des traitements médicamenteux, la **désensibilisation** (ou immunothérapie) peut être proposée. Elle est la seule méthode à visée préventive et curative de la maladie allergique. Il s'agit d'un traitement long qui vise à réduire la sensibilité de l'organisme à l'allergène.

6. Métrologie

6.1. objectifs des observatoires sanitaires

La vocation d'un réseau métrologique à visées allergologiques n'est pas de refléter toute la diversité locale (un parc, une avenue bordée de platane...) mais plutôt de traduire l'ambiance pollinique de fond, c'est à dire une situation moyenne caractéristique de la saison pollinique au niveau d'aires géographiques étendues. D'une manière générale, la concentration des grains de pollen émis par une source décroît avec la distance, ainsi plus une source de pollen est proche d'un capteur plus elle est représentée métrologiquement. Les capteurs sont donc placés préférentiellement à distance des formations végétales de manière à limiter les influences locales. Bien sur, il convient de remarquer que l'exposition aux pollens de chacun est spécifique, elle dépend du comportement individuel et des concentrations locales (qui peuvent être sensiblement différentes des concentrations de fond). Néanmoins, les caractéristiques de cette exposition spécifique présente une relation qui peut-être considérée constante avec celles de l'ambiance pollinique qui reste donc une information pertinente d'un point de vue sanitaire.

6.2. identification des grains de pollen

La morphologie du grain de pollen est caractéristique de l'espèce végétale qui le produit. La figure 1 illustre la diversité morphologique des grains de pollen à travers 4 exemples.

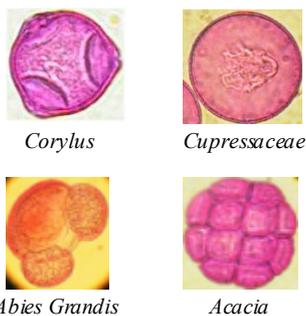


Figure 1. Illustration de la diversité morphologique des grains de pollen

L'identification des grains de pollen repose sur la taille, la forme, le nombre et la forme des ouvertures (pores et sillons) et l'architecture extrêmement variée de la membrane externe (exine). L'analyse au microscope photonique ne permet pas toujours de différencier les pollens de plantes d'espèces différentes mais appartenant à la même famille. Par exemple, dans le cas des herbacées, les déterminations sont souvent réalisées au niveau de la famille. Ainsi, les différentes "catégories" de pollen sont désignées par le terme taxon. Le nom d'un taxon désigne

donc l'espèce, le genre ou la famille de la plante qui produit le pollen selon le niveau systématique d'identification.

6.3. les capteurs

Différentes méthodes ont été développées pour réaliser des mesures du contenu pollinique de l'air. Schématiquement, les méthodes peuvent être fondées sur le principe de la sédimentation, de l'aspiration (méthode de Hirst) ou de la filtration (méthode de Cour). Les réseaux métrologiques français utilisent la méthode de Cour ou celle de Hirst.

- Méthode de Hirst (figure 2)

L'appareil de Hirst est constitué d'une pompe électrique qui prélève l'air à débit constant (10 l/min) par une buse (14 x 2 mm²) maintenue face au vent grâce à un empennage. Le capteur de Hirst a été développé dans une perspective sanitaire et le débit de prélèvement est de l'ordre

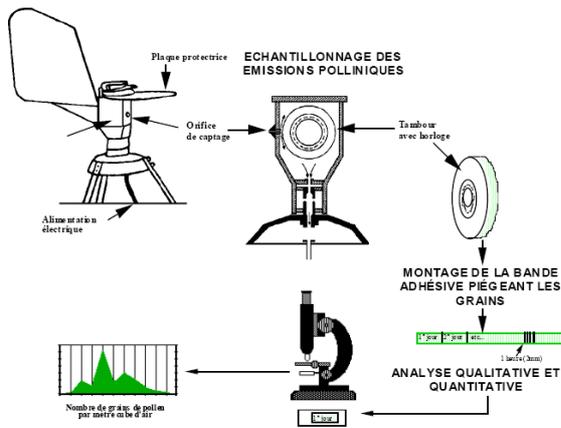


Figure 2. méthode de Hirst

du débit respiratoire de l'homme. Les particules aspirées sont piégées sur une bande adhésive fixée sur un tambour qui défile devant la buse d'aspiration grâce à un mécanisme d'horlogerie (2 mm/h). Après une semaine de fonctionnement, la bande est découpée en 7 segments qui correspondent à chaque jour de la semaine. Chaque segment est monté entre lame et lamelle dans un milieu de montage colorant et les grains sont comptés à l'aide d'un microscope photonique selon un protocole d'échantillonnage de la bande prédéfini. La concentration en pollen de l'air est calculée à partir du nombre de grains comptés en prenant en compte la méthode d'échantillonnage de la bande et le débit de prélèvement d'air.

Le choix du site d'installation des capteurs de Hirst est fait de manière à limiter les influences de la végétation locale afin de permettre la mesure de l'ambiance pollinique de fond. Les appareils sont généralement placés à une vingtaine de mètres de hauteur sur le toit d'un bâtiment de préférence en centre ville et à distance de formations végétales telles que des jardins publics. L'emplacement doit être dégagé sur 360°, sans murs ni autres obstacles susceptibles d'empêcher la libre circulation de l'air. Il est généralement admis qu'un capteur de Hirst installés selon ces recommandations fournit des données représentatives d'une zone correspondant à un disque de 20 à 30 km de rayon.

- Méthode de Cour (figure 3)

L'intercepteur pollinique de type Cour a initialement été développé dans une perspective

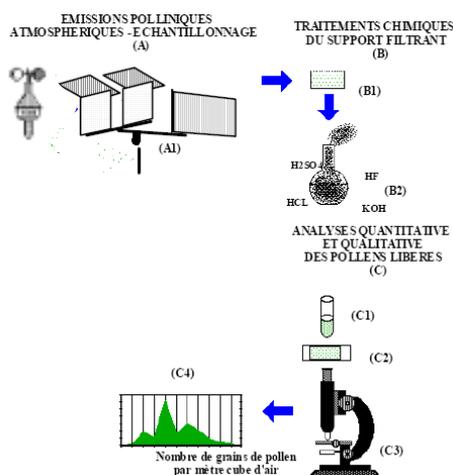


Figure 3. méthode de Cour

agronomique. Il est constitué d'une potence qui maintient deux unités filtrantes de 20 cm de côté face au vent grâce à un empennage. Les filtres sont composés de gaze hydrophile enduite d'huile de silicone. Un anémomètre totalisateur placé à proximité de l'intercepteur permet d'évaluer le volume d'air passé à travers les filtres. Il est donc possible de dériver la concentration atmosphérique de pollen à partir du nombre de grains recueillis sur le filtre pendant la période d'exposition des filtres. A la fin de la période de prélèvement, les filtres subissent un traitement visant à détruire le support filtrant ainsi que toutes les autres particules piégées en même temps que les grains de pollen. Les grains de pollen sont vidés de leur contenu cytoplasmique

par acétolyse, cela permet d'identifier les taxa avec une grande précision. La fraction mesurée de la préparation contenant les grains de pollen collectés est montée entre lame et lamelle et le comptage est réalisé à l'aide d'un microscope photonique.

Les avantages et inconvénients respectifs de la méthode de Hirst et de la méthode de Cour ont fait l'objet d'une étude spécifique commandée par la DRASS Languedoc-Roussillon intitulée "Métrologie des pollens dans l'air : étude intercomparative en région Languedoc-Roussillon" publiée en 2005. Brièvement, l'utilisation de la méthode Cour conduit à une incertitude de mesure de la concentration pollinique plus faible que celle de la méthode Hirst, en particulier à de faibles concentrations, en revanche la diminution du pas de mesure entraîne une augmentation du coût de la mise en œuvre de la méthode Cour alors que la méthode Hirst permet d'obtenir des données de concentration tous les jours voire toutes les deux heures sans modification du coût. De plus la méthode de Hirst est privilégiée dans les observatoires à visées sanitaires car c'est une méthode dynamique, sensée mieux représenter la respiration humaine.

7. Réglementation sur les pollens

La loi sur l'air et l'utilisation rationnelle de l'énergie (LAURE, loi cadre parue le 30 décembre 1996) prévoit la mise en œuvre "du droit reconnu à chacun de respirer un air qui ne nuise pas à sa santé"⁴, néanmoins il n'existe pas de réglementation spécifique concernant la mesure de pollens dans l'atmosphère malgré l'impact avéré des pollens sur la santé. Dans la région Rhône-Alpes, la publication d'arrêtés préfectoraux relatifs à l'ambrosie fait tout de même exception. Par exemple, l'arrêté préfectoral du Rhône N°2000-3261 du 20 Juillet 2000 oblige la destruction de l'ambrosie par tous les propriétaires, les locataires, les ayants droits ou quelques occupants que ce soit.

III. ETUDE DU PLAN D'ECHANTILLONNAGE DU RESEAU

1. Stratégie générale

Dans le cadre de l'étude présentée, le plan d'échantillonnage du réseau de surveillance aérobiologique est considéré optimal s'il permet d'acquérir des informations polliniques exhaustives (*i.e.* concernant l'ensemble de la zone considérée) tout en limitant les redondances (*i.e.* en mettant en œuvre le moins de capteurs possible). Le paramètre qui doit donc prévaloir à l'élaboration du plan d'échantillonnage d'un réseau est l'étendue de la zone de validité de l'information fournie par chaque capteur⁵. L'échelle à laquelle s'élabore l'information pollinique est essentiellement déterminée par deux types de facteurs : l'homogénéité du couvert végétal à l'origine des grains de pollen et les paramètres influençant le transport des grains de pollens (force et direction du vent dominant, topographie). Ainsi, l'information fournie par un capteur est représentative d'une aire géographique d'autant plus étendue qu'il est implanté dans une zone biogéographiquement homogène.

L'objectif de l'étude présentée dans cette partie est d'évaluer la représentativité du réseau métrologique du RNSA de la région Rhône-Alpes et de ses environs, et le cas échéant de faire des propositions visant à optimiser son architecture. Plus concrètement, il s'agit d'identifier les

⁴ Le terme pollution atmosphérique tel qu'il est défini dans la LAURE englobe les grains de pollen. Code de l'environnement, titre VI, article L 220-2 " Constitue une pollution atmosphérique [...] l'introduction par l'homme, directement ou indirectement, dans l'atmosphère et les espaces clos, de substances ayant des conséquences préjudiciables de nature à mettre en danger la santé humaine, à nuire aux ressources biologiques et aux écosystèmes, à influencer sur les changements climatiques, à détériorer les biens matériels, à provoquer des nuisances olfactives excessives."

⁵ zone dans laquelle l'information mesurée par un capteur représente de manière "fiable" le contenu pollinique de l'air

éventuelles redondances (*i.e.* deux sites qui fournissent des données identiques, un des deux sites pouvant donc être éliminé) ou carences du plan d'échantillonnage (*i.e.* des portions de la zone d'intérêt qui ne sont pas couvertes et qui doivent donc être équipées prioritairement lors de l'ajout de nouveaux capteurs).

Une approche possible consisterait à déterminer des zones homogènes en terme de végétation et d'exposition au vent et de s'assurer qu'un unique capteur est attribué à chacune de ces zones. Une telle approche nécessite des compétences considérables dans divers domaines tels que la botanique, la topographie, la météorologie ou la phénologie. Plus modestement, l'approche mise en œuvre est strictement statistique et repose sur les données collectées par le RNSA. La première étape de la méthode utilisée consiste à dériver un paramètre qui quantifie la "distance pollinique" séparant chaque paire de sites du réseau par comparaison des décours temporels de concentration des différents taxa. Ce paramètre est baptisé degré de proximité pollinique (DPP). Les valeurs de DPP sont ensuite traitées à l'aide d'une méthode d'analyse statistique multivariée, l'analyse de regroupement (hierarchical clustering), afin de former des groupes de sites homogènes en termes polliniques. Enfin, la manière dont se regroupent les sites est étudiée en regard de leur distribution topographique afin d'identifier des zones homogènes d'un point de vue pollinique et d'en déduire les éventuelles redondances ou carences du réseau métrologique existant.

2. Cadre de l'étude statistique

2.1. base de donnée

La base de donnée du RNSA est d'une richesse considérable. Elle contient, pour chaque station et chaque année, la valeur de concentration pollinique de chaque taxon identifiable par la méthode de Hirst (soit environ 130) à un pas de temps de 2 heures. Le réseau métrologique du RNSA compte aujourd'hui 54 stations, et certaines stations telles que celle de Lyon sont en fonctionnement depuis bientôt 20 ans. L'étude présentée se limite à 29 séries d'enregistrement qui correspondent aux données collectées de 2003 à 2005 au niveau de 9 à 11 sites selon les années.

2.2. prétraitement des données

- interpolation

Il arrive que les enregistrements soient interrompus en raison d'un défaut technique du capteur ou d'une erreur de manipulation. Dans le set de donnée considéré, 5% de la durée totale cumulée est ainsi indisponible. Seize séries de données parmi les 29 considérées présentent des enregistrements interrompus, généralement pendant une durée d'une semaine (qui correspond à la fréquence à laquelle l'analyste intervient sur le capteur pour changer la bande adhésive). La durée moyenne des interruptions est de 8 jours et leur durée maximale est de 14 jours. Les "trous" dans les données sont comblés par interpolation linéaire entre la dernière valeur mesurée avant le défaut d'enregistrement et la première valeur mesurée après celui-ci.

- calcul d'une moyenne glissante

Avant chaque analyse ou représentation graphique, les données de concentration pollinique sont moyennées sur la base d'un pas de temps de 7 jours et à l'aide d'une moyenne glissante (la valeur de chaque jour correspond à la moyenne des valeurs des 3 jours précédents et des 4 jours suivants). Ce choix a été fait car il est usuel de considérer des données polliniques hebdomadaires et parce que l'usage de la moyenne glissante permet de ne pas avoir à définir arbitrairement quel jour commence le premier "pas de temps". Ces choix peuvent bien sûr être discutés et il est à noter que la même étude pourra être réalisée rapidement à l'aide du programme matlab en modifiant ces paramètres.

2.3. approche statistique

le degré de proximité pollinique (DPP)

L'analyse statistique mise en œuvre repose sur la quantification de la "distance pollinique" qui sépare chaque paire de site. Elle peut-être définie sur la base de paramètres usuellement utilisés pour caractériser la cinétique de pollinisation tels que la concentration cumulée au cours de l'année, la date de début de pollinisation, la date de fin de pollinisation ou la durée de pollinisation. Cette approche présente l'avantage d'être simple à mettre en œuvre puisque qu'une simple différence entre deux valeurs permet de mesurer la distance entre deux sites. Néanmoins, cette approche reste insatisfaisante car :

- la définition des dates de début et de fin de pollinisation n'est pas standardisée, son choix est donc partiellement arbitraire
- ces paramètres ne caractérisent le décours temporel de concentration que de manière partielle.

Il a donc été jugé judicieux de quantifier aussi la "distance pollinique" directement sur la base des décours temporels de concentrations. Cette approche est plus lourde en terme de calcul et requiert donc l'usage d'un outil informatique adapté, mais elle présente l'avantage majeur d'être basée sur les données brutes. Les résultats présentés sont tous basés sur l'utilisation de la deuxième approche, jugée plus pertinente pour les raisons évoquées précédemment.

Un paramètre, baptisé "degrés de proximité pollinique" (DPP) a été défini afin de quantifier la distance pollinique qui sépare deux sites sur la base de la comparaison des cinétiques de concentration de pollen. Le DPP entre deux sites est calculée pour chaque taxon indépendamment. Elle est égale à un moins la somme des différences des concentrations quotidiennes mesurées par chacune des deux stations, divisée par la somme des concentrations cumulées sur l'année des deux sites. En d'autres termes, c'est le nombre de grain (par m³) émis d'une manière coïncidente sur chacun des sites, ramené au nombre total de grains (par m³) émis au cours de l'année au niveau des deux sites. La coïncidence est définie formellement par l'expression (1).

$$1 - \left(\sum_{i=1}^{i=366} |A_i - B_i| / \sum_{i=1}^{i=366} (A_i + B_i) \right) \quad (1)$$

Avec :
A_i est la concentration quotidienne de pollen au niveau du site A le jour i
B_i est la concentration quotidienne de pollen au niveau du site B le jour i.

Le DPP est un rapport qui est exprimé en pourcentage. Deux sites présentant des décours temporels de concentration d'un taxon donné identiques seraient caractérisés par un DPP de 100% pour ce taxon. La figure 4 illustre concrètement la signification du DPP.

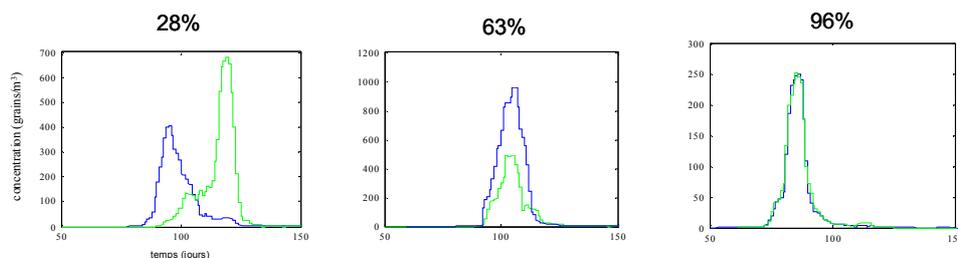


Figure 4. exemples de valeurs de degrés de proximité pollinique (DPP) liant les décours temporels de concentration pollinique de deux sites distincts (ligne bleu et ligne verte). Les valeurs de DPP sont indiquées au dessus de chaque graphique.

description de l'approche statistique

L'approche statistique présente trois étapes :

1. Les valeurs de DPP sont calculées pour chaque paire de sites, et sont stockées dans une "matrice de similarité" différente pour chaque taxon. La suite de l'analyse peut porter soit sur les valeurs de DPP calculées pour chaque taxon afin de réaliser une étude spécifique relative à un taxon, soit sur les moyennes des DPP calculées pour chaque taxon (la notation DPP_{mt} est utilisée; mt signifiant "moyenne taxon"). Cette deuxième possibilité permet de caractériser les propriétés polliniques globales du réseau et s'avère donc pertinente dans le cadre de l'étude présentée.

2. Les matrices de similarité peuvent être de dimensions considérables (nombre de sites au carré x nombre d'années), il convient donc d'utiliser une méthode de statistique multivariée pour les interpréter. L'analyse de regroupement (hierarchical clustering) a été choisie, elle vise à créer des groupes homogènes sur la base d'une matrice de similarité. Un site appartenant à un groupe (ou cluster) présente des profils de concentrations polliniques similaires à ceux des sites du même groupe et différent de ceux des sites appartenant à un autre groupe. Les résultats de l'analyse de regroupement sont représentés graphiquement sous la forme d'un dendrogramme hiérarchique. Le niveau du dendrogramme auquel se lient deux clusters représente la distance qui sépare ces clusters. Les clusters se regroupent deux à deux jusqu'à ce que tous les sites soient inclus dans le dendrogramme. Ultiment chaque site est lié aux autres à un certain niveau dans le dendrogramme. Afin d'estimer à quel point le dendrogramme est un reflet fiable de la distance qui lie effectivement chaque paire de site, le coefficient de corrélation cophénétique est calculé. Il dérive de la comparaison entre les distances polliniques réelles (celles contenues dans la matrice de similarité) et les distances cophénétiques, c'est à dire celles représentées dans le dendrogramme.

3. Le dendrogramme représentant le regroupement des sites sur la base de leurs propriétés polliniques est étudié en relation avec la distribution topographique des sites afin de délimiter des zones homogènes en termes polliniques. Il est ainsi possible d'identifier les redondances ou carences éventuelles du réseau.

2.4. définition des limites de l'étude

L'étude présentée se limite à 13 taxa: l'ambrosie (*Ambrosia*), le bouleau (*Betula*). Le cyprès (*Cupressaceae*), les graminées (*Poaceae*), le platane (*Platanus*), l'aulne (*Alnus*), le chêne (*Quercus*), le peuplier (*Populus*), le châtaigner (*Castanea*), le frêne (*Fraxinus*), le noisetier (*Corylus*), l'oseille (*Rumex*) et le saule (*Salix*). Ils ont été sélectionnés sur la base du critère de "risque allergique" défini par le RNSA : ces taxa ont tous présenté au cours de l'année 2005 un "risque allergique" de niveau moyen, élevé ou très élevé au moins une semaine dans l'année. S'il n'existe pas de consensus sur la définition et la validité des niveaux de "risque allergique", il est tout de même admis que les taxa sélectionnés constituent une préoccupation d'ordre sanitaire dans la région Rhône-Alpes.

Deux études ont été menées afin de caractériser le réseau métrologique:

- **L'étude 1** porte uniquement sur l'année **2005**, elle inclut **13 taxa** et concerne **11 stations**. Le réseau en compte 16 mais 3 d'entre elles ont pour vocation la métrologie exclusive du pollen d'ambrosie. Les stations de mesure concernées par l'étude sont représentées figure 5 (points pleins bleus et rouges).

- **L'étude 2** porte sur les années **2003 à 2005** et sur les **9 stations** qui sont à l'origine de données pour chacune des 3 années de la période d'intérêt. La figure 5 présente la position topographique de ces sites (points pleins bleus). Seuls **9 taxa** sont considérés car des stations telles que celles de Lyon ou de Roussillon présentent des enregistrements ayant débutés trop tard certaines années pour mesurer les pollens "précoces" (cyprès, aulne et noisetier). De plus

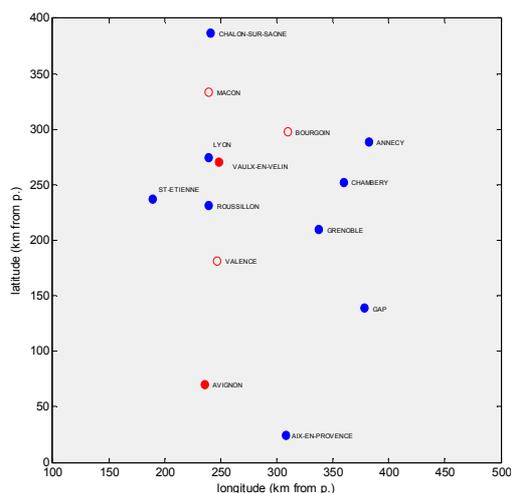


Figure 5. localisations des stations de mesure du RNSA. En bleu : stations opérationnelles de 2003 à 2005; en rouge : les stations implantées en 2005; en rouge vide : les stations dévolues exclusivement à la météorologie de l'ambrosie

3. Résultat – Discussion

3.1. étude 1 : année 2005

- exemple d'un taxon individuel : l'oseille (*Rumex*)

Les sites s'organisent en clusters distincts sur la base de leurs propriétés polliniques

En préambule de la présentation des résultats portant sur les 13 taxa, l'analyse relative à un taxon individuel est exposée, cela permet d'illustrer la méthodologie et d'introduire quelques principes régissant l'organisation de l'information pollinique. La figure 6 présente les décours temporels de concentration de grain de pollen d'oseille (*Rumex*) enregistrés au niveau des 11 stations de la région d'intérêt. Elle permet d'apprécier subjectivement le degré de similitude entre les enregistrements réalisés au niveau des différents sites de la région d'intérêt. En premier lieu, il apparaît que les sites présentent une grande diversité de profils polliniques (Cf. par exemple les enregistrements de Gap et de St Etienne). Néanmoins, au sein de cette diversité émergent des groupes de sites partageant les mêmes propriétés. Par exemple, il est possible de noter que les profils de concentrations d'Annecy, Chambéry, Chalon-sur-Saône et Lyon sont similaires. De la même manière, Aix-en-Provence et Avignon présentent aussi des caractéristiques polliniques communes. En revanche Gap semble avoir un profil spécifique.

Une approche statistique est nécessaire pour caractériser objectivement le degré de similitude entre les enregistrements des différents sites. La matrice de DPP de l'oseille n'est pas figurée, mais à titre d'exemple, les valeurs minimale et maximale sont présentées respectivement par les paires de sites Gap-St Etienne (13%) et Annecy-Chambéry (88%). La méthode d'analyse de regroupement (hierarchical clustering) permet d'exploiter la matrice de DPP de manière statistique en groupant les sites qui présentent des décours temporels semblables. Le résultat est représenté sous la forme d'un dendrogramme hiérarchique (figure 7). Le coefficient de corrélation cophénétique est de 0.9, cela indique que le dendrogramme est un reflet fiable des données brutes de DPP. Différents clusters peuvent être identifiés. Un premier cluster (ligne rouge) est composé de 4 sites (Annecy, Chambéry, Chalon-sur-Saône et Lyon) aux propriétés polliniques très proches puisque la valeur de DPP minimale au sein du groupe est de 76%. Roussillon et St Etienne forment un groupe homogène, de même que Vaulx-en-Velin et

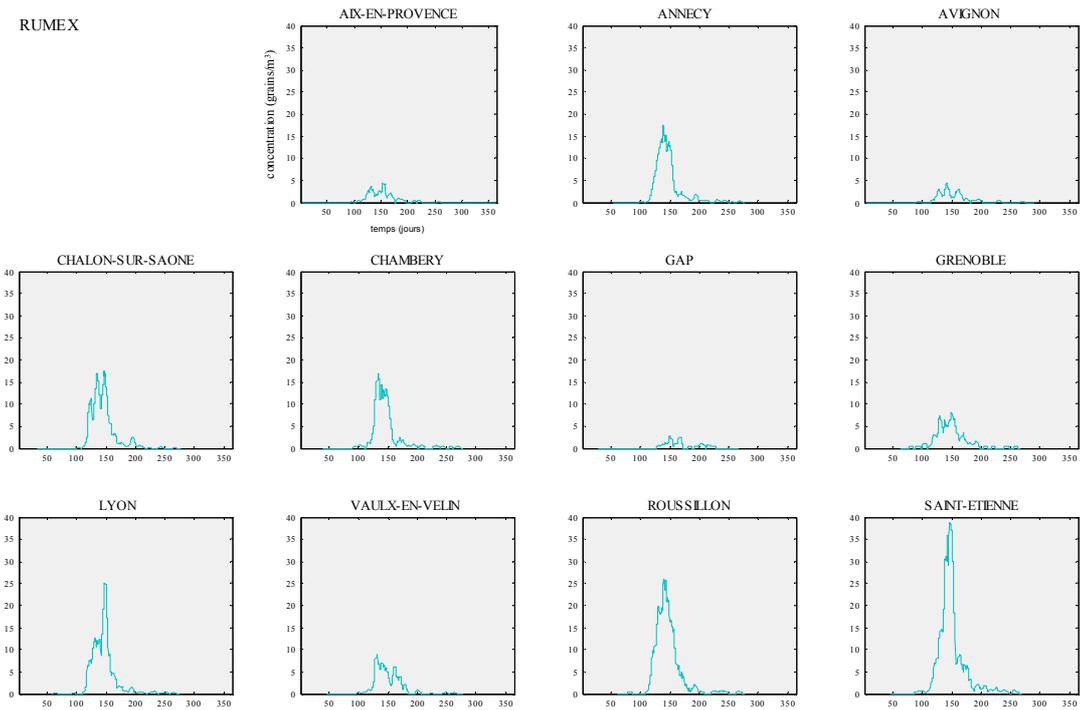


figure 6. Décours temporels de la concentration de *Rumex* (oseille) au niveau des onze sites étudiés

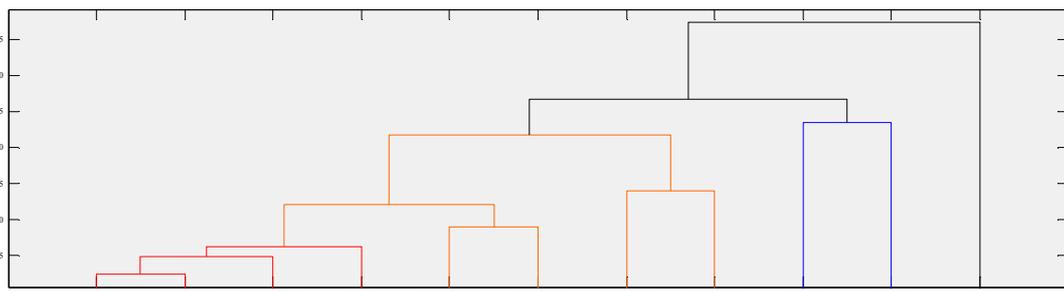


figure 7. Dendrogramme hiérarchique représentant les différents clusters homogènes

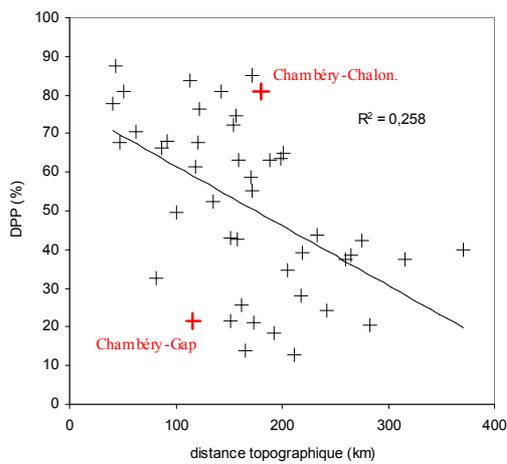


figure 8. Relation entre la distance 'pollinique' et la distance topographique

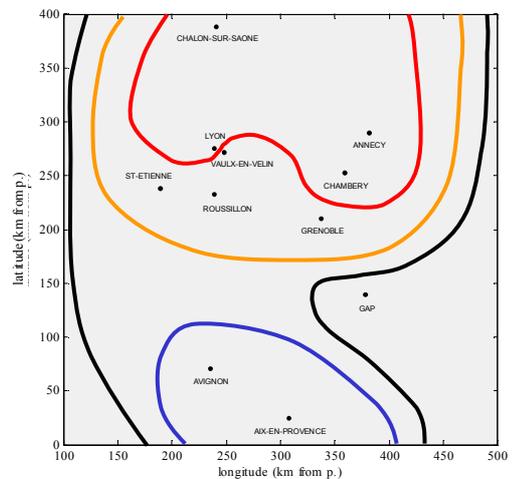


figure 9. Représentation des zones d'homogénéité pollinique

Grenoble. Ces deux clusters se regroupent avec le premier pour former un cluster plus vaste (ligne orange), mais évidemment moins homogène (la valeur minimale de DPP est atteinte par la paire Grenoble-St Etienne et n'est que de 43%). Ce cluster se regroupe avec un cluster formé de la paire Aix-en-Provence-Avignon (DPP = 66%). Enfin, Gap rejoint l'ensemble des autres sites pour former le dendrogramme complet. Cela témoigne de la spécificité pollinique de Gap.

la région Rhône-Alpes présente 3 zones d'homogénéité pollinique

Afin de comprendre les principes qui régissent l'organisation topographique de l'information pollinique les résultats de l'analyse de regroupement sont transférés dans un référentiel topographique. En première approche, il est possible de considérer la relation entre la distance pollinique (DPP) et la distance topographique (figure 8). Il semble que si, d'une manière générale, la DPP diminue avec l'augmentation de la distance topographique, la relation entre ces deux paramètres est faiblement corrélée ($R^2=0.3$). Cela, indique que le DPP n'est pas strictement déterminé par la distance topographique et qu'il pourrait exister dans la zone d'étude des régions caractérisées par des propriétés bioclimatiques différentes. A titre d'exemple, il est intéressant de considérer les paires de sites Chambéry-Gap et Chambéry-Chalon. En effet, alors que la distance qui sépare Chambéry de Gap est sensiblement la même que celle qui sépare Chambéry de Chalon-sur-Saône, ces deux paires de sites présentent des DPP considérablement différents. Ainsi, Chambéry et Chalon-sur-Saône semblent appartenir à une zone polliniquement homogène alors que Chambéry et Gap pourraient dépendre de régions "polliniques" différentes.

Afin d'étudier plus systématiquement, la manière dont s'élabore l'information pollinique dans la zone étudiée, la distribution topographique des stations de mesure et le groupement des sites en clusters polliniquement homogènes sont représentées conjointement (figure 9). Les limites des clusters sont figurées par des lignes de couleur, le code de couleur est le même que celui utilisé figure 4 (les limites des clusters sont placées de façon à inclure les sites appartenant à un cluster donné, mais leur position précise est arbitraire). Sur la base des données considérées, il apparaît, et cela n'a rien d'étonnant, que les sites formant des groupes homogènes d'un point de vue pollinique sont des sites voisins. Plus spécifiquement, il est possible d'identifier des zones topographiques homogènes polliniquement. Par exemple, les sites de Chalon-sur-Saône, Lyon, Annecy et Chambéry appartiennent à une zone d'une forte homogénéité pollinique, puisque tous ces sites présentent une coïncidence d'au moins 76% avec chacun des autres. Les limites nord, est et ouest de la zone étudiée ne permet pas de déterminer les dimensions de la zone d'homogénéité, néanmoins elle couvre au moins l'aire incluse entre les quatre sites (soit environ 10^4 km²; à moins d'une spécificité locale non échantillonnée par les 4 stations). Les sites provençaux d'Aix-en-Provence et Avignon forment aussi une zone homogène en terme de cinétique de concentration d'oseille. Ces sites se trouvent en bordure de la zone d'intérêt et la dimension de la zone homogène ne peut donc pas être évaluée. Gap est le site de la région qui présente la plus grande spécificité pollinique, cela tient vraisemblablement à une relative spécificité de la végétation et du relief environnant. Il convient d'observer que les sites de Lyon et Vaulx-en-Velin, en dépit de leur grande proximité topographique (les sites sont séparés de 6 km), présentent des cinétiques de concentrations polliniques sensiblement différentes (DPP=59%) et appartiennent donc à des clusters différents. Cette incongruité de l'organisation topographique et de l'information pollinique est constante pour tous les taxa et sera discutée dans le paragraphe consacré à l'étude conjointe des 13 taxa.

Il apparaît donc qu'en ce qui concerne l'oseille, l'information pollinique s'élabore à une échelle plus petite que celle de la région étudiée. Il est possible d'identifier des zones d'homogénéité

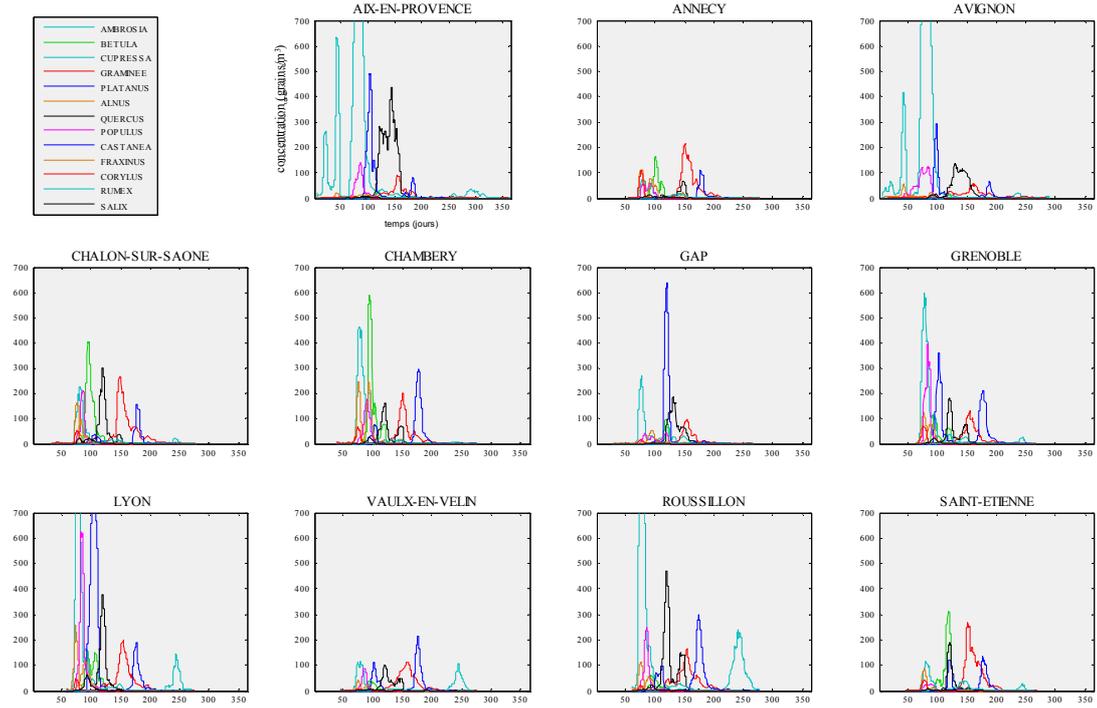


figure 10. Décours temporels de la concentration de 13 taxa au niveau des onze sites étudiés

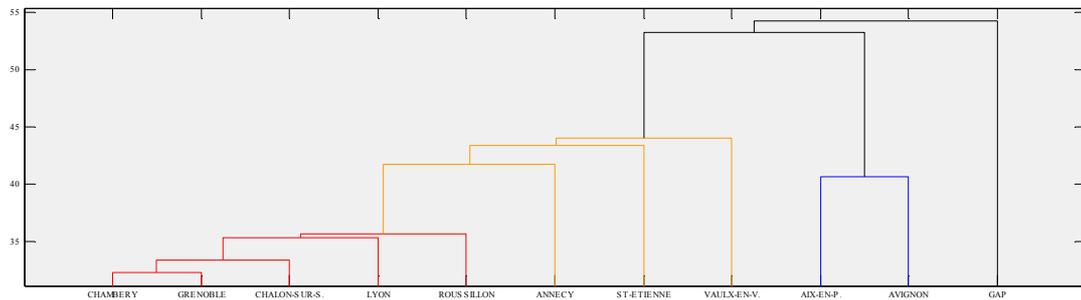


figure 11. Dendrogramme hiérarchique représentant les différents clusters homogènes

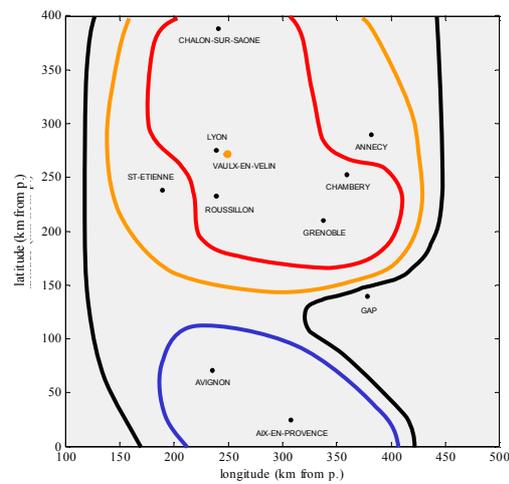


figure 12. Représentation des zones d'homogénéité pollinique

pollinique (dans la limite du plan d'échantillonnage du réseau) dont la dimension variable peut être interprétée en terme d'homogénéité biogéographique. Les implications qu'ont ces observations sur l'architecture du réseau ne sont pas discutées ici. En effet, la vocation du réseau n'est pas de mesurer l'oseille mais de mesurer tous les taxa, l'évaluation de son plan d'échantillonnage doit donc être réalisée sur la base de tous les taxa. C'est l'objet des paragraphes suivants.

- **Analyse portant sur les 13 taxa sélectionnés**

L'étude ne peut pas rendre compte de la spécificité de chaque taxon

L'analyse décrite dans le cas de l'oseille a été réalisée pour chacun des 13 taxa sélectionnés (résultats non présentés). La comparaison des résultats relatifs à chaque taxon permet de dégager des principes généraux d'organisation topographique de l'information pollinique communs à presque tous les taxa : Gap est un site présentant des propriétés polliniques très spécifiques, Aix-en-Provence et Avignon, constitue une zone provençale homogène et les autres sites constituent une vaste zone relativement homogène. Néanmoins, au-delà de ces règles générales, le regroupement des sites en clusters homogènes se fait de manière relativement spécifique pour chaque taxon. Ainsi, les résultats qui vont être présentés doivent être considérés comme représentatifs d'une situation moyenne, valables pour la plupart des taxa, mais qui ne peuvent pas représenter la spécificité de chaque taxon. Par exemple, l'ambrosie dont la distribution topographique est particulière, est mal représentée par le résultat global.

Les sites se regroupent en trois clusters majeurs

La figure 10 présente les décours temporels de concentration des 13 taxa enregistrés au niveau des 11 sites de la région d'intérêt. La quantité d'information représentée graphiquement ne permet pas une analyse détaillée de la figure, mais il est néanmoins possible d'observer que la région présente une grande diversité pollinique et qu'il n'apparaît pas de sites aux profils manifestement semblables. A titre d'exemple, les paires de sites présentant les plus grandes valeurs de DPP moyennées pour tous les taxa (DPP_{mt}) sont Chambéry-Grenoble (68%), Chambéry-Chalon-sur-Saône (67%) et Grenoble-Lyon (65%).

La manière dont se regroupent les sites sur la base du décours temporel de tous les taxa est représentée par un dendrogramme hiérarchique (figure 11; coefficient de corrélation cophénétique = 0.9). Un premier cluster homogène regroupant Chambéry, Grenoble, Chalon-sur-Saône, Lyon et Roussillon peut-être identifié (ligne rouge). A l'intérieur de ce cluster les valeurs de DPP_{mt} sont comprise entre 57% (Chambéry-Roussillon) et 68% (Chambéry-Grenoble). Ce cluster peut être étendu en incluant Annecy, St-Etienne et Vaulx-en-Velin (ligne orange), mais cette extension altère considérablement l'homogénéité du groupe. Un deuxième groupe homogène constitué d'Aix-en-Provence et Avignon (DPP_{mt}=59%, ligne bleue) rejoint ensuite le premier cluster. Néanmoins, le "cluster bleu" est polliniquement très distant du "cluster orange" puisqu'il partage au mieux une DPP de 45% (Grenoble-Aix-en-Provence). Le site de Gap ne participe à aucun groupe avant de rejoindre le cluster total (ligne noire).

Les limites des groupes polliniquement homogènes sont représentées sur une carte figurant la position des stations de mesure (figure 12). D'une manière générale, les zones polliniquement homogènes s'organisent de manière harmonieuse sur la carte, c'est à dire que chaque cluster est constitué de sites voisins et que l'extension d'un cluster (par exemple l'extension du "cluster de rouge" au "cluster orange") se traduit par une extension de la zone géographique.

Vaulx-en-Velin : représentativité entachée par l'influence de la végétation locale

La station de Vaulx-en-Velin constitue une exception frappante au principe d'organisation de l'information pollinique décrit précédemment. En effet, alors que cette station se situe à 6 km de celle de Lyon et qu'elle est incluse à l'intérieur d'une zone délimitée par les sites formant le "cluster rouge", elle est, d'un point de vue pollinique, relativement éloigné de ces sites (Vaulx-en-Velin ne rejoint le "cluster rouge" qu'après Annecy et St-Etienne pour former le "cluster orange"). La raison de l'incohérence entre la position topographique et les propriétés polliniques de la station de Vaulx-en-Velin tient vraisemblablement au choix de son site d'installation. En effet, les capteurs doivent être placés de manière à limiter les influences de la végétation locale afin de permettre la mesure de l'ambiance pollinique de fond. Les appareils doivent donc être installés à une vingtaine de mètres de hauteur sur un site dégagé sur 360° (e.g. sur le toit d'un bâtiment), de préférence en centre ville. Dans le cas du capteur de Vaulx-en-velin, ces recommandations ne sont pas respectées puisqu'il est placé à environ 5 mètres de hauteur et à proximité de champs. Les informations fournies par ce capteur reflètent donc le contenu pollinique local et il ne doit donc pas être inclus dans un réseau dont la vocation est de mesurer la concentration pollinique de fond. Les 10 autres capteurs de la portion de réseau étudiée sont généralement implantés de manière appropriée, mais la difficulté pratique de trouver un bâtiment adapté introduit une certaine variabilité dans la hauteur des capteurs (max=50 m, min=12 m, moyenne=27 m, écart type=12 m). Les résultats présentés sont vraisemblablement affectés par ce défaut de standardisation des hauteurs d'implantation des capteurs.

Trois zones d'homogénéité pollinique majeures peuvent être identifiées en Rhône-Alpes

Les stations du réseau de la région d'intérêt (à l'exception de Vaulx-en-Velin) permettent de délimiter des zones homogènes d'un point de vue pollinique. La région incluant les sites appartenant au "cluster rouge" est relativement homogène puisque (DPP_mt minimale = 57%). Cette zone est caractérisée par des concentrations relativement élevées de pollens de châtaigner, peuplier, graminées et cyprès. Il faut noter tout de même que les pollens de bouleau et d'ambrosie présentent des propriétés assez variables au sein de la zone incluant les sites du "cluster rouge." L'étendue de cette zone est impossible à estimer car elle situe en bordure de la zone d'intérêt. Les sites de St-Etienne et Annecy se regroupent avec ce cluster pour former un cluster moins homogène. Ces sites se situent en bordure de la zone étudiée et il est possible que si d'autres stations avaient été incluses dans l'étude, ces sites se seraient regroupés avec d'autres (plus au nord et à l'est dans le cas d'Annecy et plus à l'ouest dans le cas de St-Etienne) pour former des groupes aux propriétés polliniques proches. Les stations d'Avignon et d'Aix-en-Provence forment un groupe polliniquement distinct, caractérisé par une très forte représentation des pollens de cyprès, et à l'exception des pollens de platane, de chêne et de peuplier, des concentrations polliniques relativement faibles. L'étendu géographique de la zone homogène qu'ils représentent ne peut être définie sur la base du réseau étudié. Gap est le site qui présente la plus grande spécificité pollinique puisqu'il regroupe avec les autres sites en dernier. Cela peut s'expliquer par le fait que ce site est un peu excentré de la zone étudiée, néanmoins il est vraisemblable que Gap appartienne à une zone polliniquement homogène de petite taille en lien avec la spécificité locale de la végétation et du relief.

Réflexion sur l'architecture du réseau métrologique

Il est donc possible d'identifier différentes zones d'homogénéité pollinique (dans la limite des points échantillonnés) dont les limites peuvent être interprétées en terme d'homogénéité biogéographique. Ces observations peuvent servir de base à une réflexion relative à l'architecture du réseau métrologique. Deux questions se posent dans le cadre de l'optimisation d'un réseau métrologique : Si un site devait être ajouté, où devrait-il l'être ? Si une station de mesure devait être enlevée, laquelle faudrait-il enlever et quelle conséquence cela aurait-il sur

la qualité de l'information fournie par le réseau métrologique ? Le dendrogramme montre qu'il existe trois clusters distants, le cluster orange, le cluster bleu et le cluster composé du seul site de Gap (figure 7). La représentation des limites de ces clusters sur une carte topographique indique que ces clusters forment des zones géographiques distinctes (figure 8). Le principe de construction des clusters est tel que les propriétés polliniques sont homogènes au sein des clusters et différentes entre les clusters. C'est donc "entre" les clusters que les propriétés polliniques sont mal caractérisées et c'est là qu'il faut implanter prioritairement des sites. De manière plus concrète, la région située au sud d'une ligne liant Grenoble à St Etienne et au nord d'Avignon est la zone la plus mal "représentée" par le réseau existant. Si une nouvelle station devait être implantée, Valence constituerait un choix rationnel en terme statistique.

En revanche, si le nombre de stations doit être réduit, il faut considérer la règle selon laquelle l'élimination d'un site affecte d'autant moins la représentativité du réseau que ce site se trouve dans une zone d'une grande homogénéité pollinique. Dans le cas théorique où deux sites sont strictement redondants, il est possible d'éliminer un site et de considérer les informations collectées par le deuxième comme valide pour les deux sites. En pratique, aucune paire de site de la portion de réseau considérée ne présente de DPPmt égal à 100%. Cela implique, que dans le cadre d'une stratégie purement métrologique, l'élimination d'un capteur implique nécessairement une dégradation de l'information fournie par le réseau. La question qui demeure est : à quel point cette information est-elle dégradée quand le contenu pollinique de l'air mesuré au niveau d'une station est directement extrapolé à un autre site ? Dans la zone d'étude, le cluster de plus grande homogénéité est constitué par les deux sites voisins Chambéry et Grenoble (DPPmt=68%). Ainsi, si un capteur devait être éliminé, ce serait prioritairement l'un de ces deux là. La figure 13 présente, pour ces deux sites, le décours temporel de la concentration de chaque taxon et permet d'apprécier subjectivement la proximité pollinique de ces deux sites. La valeur de DPP est supérieure à 70% pour la majorité des taxa (e.g. chataigner DPP=88%), néanmoins 5 taxa ont une valeur de DPP inférieure à 70%, c'est le pollen de platane qui présente la plus basse valeur (DPP=36%). La question est donc la suivante : ces deux sites fournissent-ils une information suffisamment semblable pour permettre aux utilisateurs des données fournies par le réseau de faire leur travail dans de bonnes conditions si l'une des stations était éliminée ? Cette question ne relève pas de la statistique mais de l'allergologie.

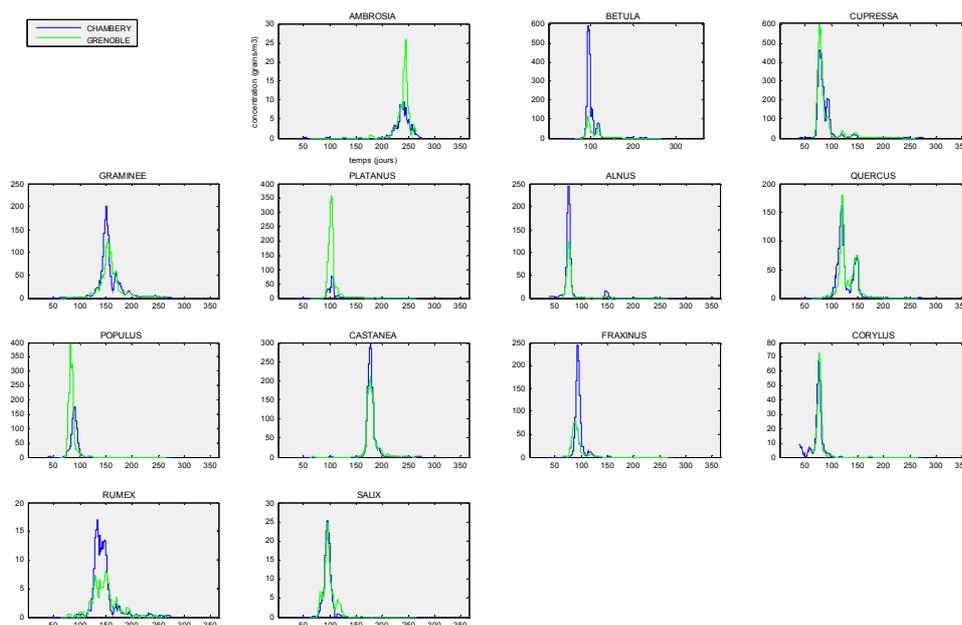


Figure 13. Décours temporel de chaque taxon au niveau des sites de Grenoble et Chambéry. Ces sites présentent un DPPmt de 68%. Les valeurs de DPP de chaque taxon sont les suivantes : ambrosia 70%, betula: 43%, cupressa 84%, graminee 77%, platanus 36%, alnus 70%, quercus 86%, populus 42%, castanea 88%, fraxinus 50%, coryllus 84%, rumex 68% ,salix 84%.

Certains capteurs pourraient subir une influence locale ?

Le cas de Vaulx-en-Verin est un exemple frappant de l'influence de la végétation locale sur les enregistrements d'un capteur, mais l'analyse indépendante de chaque taxon semble indiquer que d'autres (rares) capteurs subissent une telle influence. En effet, dans certains cas le regroupement en clusters et la distribution topographique des sites paraît incohérent en regard des règles générales dégagées dans l'étude (résultats non présentés). Par exemple, en ce qui concerne le platane, le site de Roussillon se regroupe avec Lyon à un haut niveau dans le dendrogramme et, bien que cette hypothèse relève en partie d'un raisonnement circulaire, il est possible que cela soit le reflet de l'environnement local d'un des capteurs.

L'analyse statistique mise en œuvre permet de fournir des indications rationnelles quant à la position des sites à ajouter ou à enlever prioritairement, mais la décision d'ajouter ou d'enlever des stations doit se baser sur les besoins des allergologues et sur des critères économiques. De plus l'approche statistique vise à élaborer un réseau qui représente toute la région mais ne considère pas la distribution de la population sur le territoire. Ces conclusions sont fondées sur l'étude de l'année 2005, il convient de s'assurer de la stabilité des propriétés du réseau au cours des années précédentes.

3.2. étude 2 : période 2003-2005

L'étude présentée dans les paragraphes suivants, porte sur les années 2003 à 2005, 9 sites et 9 taxa. Le nombre de sites et de taxa est inférieur à ceux considérés dans l'étude précédente car le réseau et les périodes de mesure ont changé au cours de ces 3 années. Les 2 sites qui ne sont pas inclus dans l'analyse sont : Vaulx-en-Verin, pour les raisons de localisation évoquées précédemment, et Avignon dont le capteur ne fournit des données concernant toute la saison pollinique qu'à partir de 2004. L'approche statistique est identique à celle mise en œuvre dans l'étude précédente. La figure 14 présente le dendrogramme hiérarchique qui illustre le regroupement en clusters homogènes des différents sites au cours des 3 années.

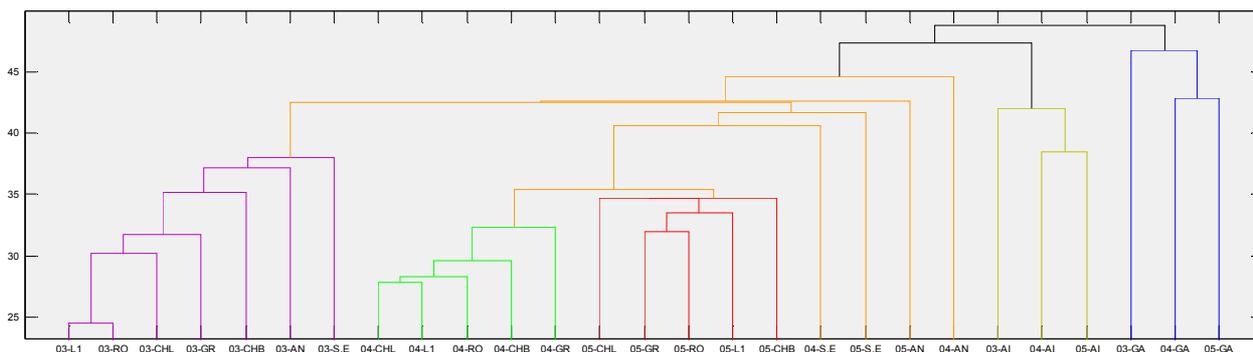


Figure 14. Dendrogramme hiérarchique illustrant le regroupement en cluster de 9 sites au cours de 3 années. Les abréviations utilisées sont les suivantes : 03, 2003; 04, 2004, 05, 2005; AI, Aix-en-Provence; AN, Annecy; CHL, Chalon-sur-Saône; CHB, Chambéry; GA, Gap; GR, Grenoble; L1, Lyon; RO, Roussillon; S.E., St Etienne.

D'une manière générale deux types de clusters peuvent être identifiés, ceux qui regroupent des enregistrements d'un même site mais réalisés des années différentes et ceux qui regroupent des enregistrements de différents sites mais d'une même année. Les clusters du premier type

sont au nombre de deux et concernent les sites de Gap (ligne bleue) et d'Aix-en-Provence (ligne jaune). Cela est cohérent avec les résultats de la première étude puisque ces sites se trouvaient déjà regroupés dans des clusters distincts des autres sites (Aix-en-Provence était regroupé avec Avignon mais ce site est absent de l'étude en cours). Les trois clusters du deuxième type ("cluster violet", "vert" et "rouge") regroupent les sites de Chalon-sur-Saône, Lyon, Roussillon, Chambéry et Grenoble, c'est à dire les 5 sites déjà fortement liés dans la première étude. Les enregistrements de St Etienne et Annecy se joignent à ces clusters à un plus haut niveau conformément à ce qui avait été constaté dans la première étude.

La structure de base des clusters est stable d'une année à l'autre

L'analyse des enregistrements des trois années consécutives montre que les propriétés essentielles de l'organisation topographique de l'information pollinique sont stables dans le temps. Cela n'a rien d'étonnant puisque les facteurs qui déterminent les limites des zones d'homogénéité, homogénéité du couvert végétal et direction et la force du vent, sont relativement stables d'une année à l'autre. Les conclusions de l'étude 1 relatives à d'éventuelles modifications du réseau ont donc une validité qui dépasse la seule année 2005. Néanmoins, parce que l'étude 1 porte sur 4 taxa de plus que l'étude 2, la manière dont se regroupent les enregistrements de 2005 des sites de Chalon-sur-Saône, Grenoble, Roussillon, Lyon et Chambéry est sensiblement différente dans les deux études. Cela implique notamment que la paire de site qui présente la plus haute DPP_mt au cours des 3 années est Lyon-Roussillon (2003 : DPP_mt=75%; 2004 : DPP_mt=72%; 2005 : DPP_mt=61%) et non pas Chambéry-Grenoble comme c'était le cas dans l'étude 1. La valeur moyenne de DMS_mt sur les 3 années est de 69%, la discussion relative à la validité de l'élimination d'un de ces deux sites est la même que celle menée dans le cadre de l'étude 1.

La variabilité interannuelle au niveau d'un site donné est élevée

Une caractéristique essentielle du regroupement est que les sites qui appartiennent à une même zone d'homogénéité forment des clusters distincts pour chaque année, en d'autres termes la variabilité interannuelle du contenu pollinique de l'air pour un site donné est supérieure à la variabilité entre les sites voisins pour une année donnée. La production, la libération et le transport des grains de pollen déterminent la concentration en grain de pollen à un endroit et à un instant donné. Ces trois paramètres sont influencés par des facteurs météorologiques, donc la variabilité interannuelle des conditions météorologiques peut expliquer la forte variabilité interannuelle des enregistrements polliniques. A titre d'exemple, l'année 2003 a été exceptionnelle en termes météorologiques et c'est en effet cette année qui présente la plus grande spécificité pollinique. Cela se traduit au niveau du dendrogramme (figure 10) : les enregistrements de 2003 de Gap et Aix-en-Provence rejoignent leur cluster respectifs à un haut niveau et les enregistrements d'Annecy et St Etienne se regroupent avec le cluster "2003" ("cluster violet") à bas niveau contrairement aux enregistrements de ces mêmes sites réalisés en 2004 ou 2005. Dans le cadre de l'étude de l'optimisation du fonctionnement météorologique, ce résultat conduit à conclure que si le nombre d'enregistrement devait être réduit, cela devrait se traduire par l'élimination d'une station plutôt que des enregistrements limités à certaines années. En outre, comme cela sera exposé en détail dans la partie III, ce résultat a de fortes implications en ce qui concerne le choix d'une stratégie de modélisation.

4. Conclusions – possibilités d'optimisations

conclusions de l'étude

Les études réalisées ont permis de caractériser les règles qui régissent l'élaboration de l'information pollinique au sein de la région Rhône-Alpes et d'évaluer la stabilité de ces règles

d'une année à l'autre. Les possibilités d'optimisation du réseau métrologique de RNSA ont été explorées sur la base des résultats de ces études.

- Trois zones homogènes d'un point de vue pollinique peuvent être identifiées dans la région d'intérêt :

Une zone de grande homogénéité délimitée par les sites de Chalon-sur-Saône, Lyon, Roussillon, Grenoble et Chambéry. Cette zone peut-être étendue à Annecy et St Etienne mais cela implique une dégradation de l'homogénéité pollinique.

Une zone homogène regroupant les deux sites provençaux : Avignon et Aix-en-Provence

Le site de Gap présente une grande spécificité pollinique, vraisemblablement liée aux caractéristiques locales de la végétation et du relief.

- L'étude de la position des capteurs en relation avec leur regroupement en clusters permet d'identifier la zone située au sud de Grenoble et au nord d'Avignon comme la région la moins bien "représentée" par le réseau existant. Ainsi, Valence est potentiellement le site où l'implantation d'un nouveau capteur augmenterait le plus la représentativité du réseau.

- Les deux sites présentant le plus grand degré de proximité pollinique sont selon les paramètres de l'étude : Chambéry et Grenoble (étude 1 : année 2005, 13 taxa; DPP_mt=68%) ou Lyon et Roussillon (étude 2 : années 2003-2005, 9 taxa; DPP_mt moyenne sur les 3 années: 69%). Si une station devait être éliminée, c'est un site parmi ces deux paires de sites qui devrait être choisi et il faudrait alors considérer que les informations fournies par un site peuvent être directement transposées à l'autre site. Est-ce que 69% de similitude est une valeur suffisante pour permettre aux utilisateurs des données du réseau de remplir leur mission ? Cette question ne relève pas de la statistique mais de l'allergologie et de l'aérobiologie, elle nécessite donc d'être discutée de manière collégiale. Néanmoins, en première approximation, ce résultat indique que le réseau étudié ne présente pas de redondance stricte.

- La variabilité interannuelle de l'information pollinique fournie par un site donné est plus grande que la variabilité entre les enregistrements réalisés la même année par des capteurs appartenant à une zone homogène. Cela n'est pas étonnant car les facteurs qui déterminent les zones polliniquement homogènes (la végétation et l'exposition au vent) sont stables d'une année à l'autre, alors que les paramètres influençant le contenu pollinique de l'air (certains facteurs météorologiques comme la température ou l'ensoleillement) présentent une certaine variabilité interannuelle. Ce résultat implique que s'il est nécessaire de réduire le nombre d'enregistrements, cela doit se traduire par la réduction du nombre de capteurs (Cf point précédent) plutôt que par la mesure de tous les sites une année sur deux.

Si les études présentées permettent d'identifier les sites les plus opportuns à l'implantation d'un nouveau capteur ou les stations dont l'élimination affecterait le moins la représentativité du réseau, la décision d'ajouter ou d'enlever un capteur est conditionnée par des facteurs sanitaires et économiques. De plus la répartition de la population, paramètre occultée dans l'approche présentée, doit être considérée afin d'évaluer la représentativité du réseau, non pas en terme purement géographique, mais sur la base de la proportion de la population habitant dans une région où le contenu pollinique de l'air est mesuré.

Remarques générales

Les données météorologiques fournies par le RNSA sont d'une richesse considérable, néanmoins il convient de souligner certaines de leurs limites :

- Les enregistrements présentent parfois des données manquantes, et comme cela a été décrit dans la partie méthode, 5% de la durée totale cumulée est ainsi indisponible dans le set de données étudié. Cela porte préjudice tant à la qualité des études du type de celles présentées dans ce rapport, qu'au travail du RNSA qui consiste à produire l'indice pollinique hebdomadaire de chaque région. Il est donc nécessaire de limiter les défauts d'enregistrements, mais aussi de valider et d'optimiser la méthode d'interpolation utilisée dans les études présentées.

- Les enregistrements débutent et finissent à des dates différentes selon les sites, cela rend délicat les études statistiques à grande échelle. Une standardisation des périodes d'enregistrement serait donc souhaitable. Le choix des limites de la période d'enregistrement relève de l'aérobiologie et de l'allergologie, mais il est d'ores et déjà possible de noter qu'un site comme Aix-en-Provence présente un contenu aérobiologique significatif pratiquement tout au long de l'année.

- Le choix du site d'implantation d'un capteur détermine sa représentativité. Le cas du capteur de Vaulx-en-Verin illustre qu'un choix inadapté peut amener à devoir rejeter une partie des données collectées. Il semble, mais cela nécessiterait une étude spécifique, que de rares capteurs subissent l'influence de taxa locaux rendant l'exploitation des enregistrements relatifs à ces taxa délicates. Dans ce contexte là, il serait intéressant de réaliser une étude préliminaire à l'implantation d'un capteur qui viserait à évaluer l'influence de la végétation locale au niveau du site sélectionné. Le RNSA pratique une étude du couvert végétal environnant avant d'installer un capteur, mais en complément il pourrait être envisagé d'utiliser une station temporaire afin d'estimer la représentativité du site sur la base d'une étude du type de celle présentée dans ce rapport.

- Une caractéristique essentielle de la stratégie météorologique du RNSA est d'être fondée exclusivement sur un réseau de capteur de Hirst. L'utilisation d'intercepteurs polliniques de Cour conduit à une incertitude de mesure de la concentration pollinique plus faible que celle de la méthode Hirst. Cette propriété est essentielle puisque le début de la saison pollinique doit être détecté le plus précocement possible afin que les patients allergiques puissent commencer leur traitement avant le déclenchement des symptômes. Néanmoins, la lourdeur de cette méthode proscrit son usage à grande échelle. Les deux méthodes météorologiques présentent des avantages complémentaires et un réseau mixte fondé sur les deux types de capteurs pourrait être envisagé. En pratique, la situation française actuelle n'est pas très différente puisque des réseaux de capteurs de Cour cohabitent avec le réseau du RNSA. Toutefois ces observatoires ne sont pas gérés par les mêmes organismes, cela exige de faire appel à des interlocuteurs différents, qui utilisent des conventions scientifiques différentes, pour accéder aux informations (base de données ou information au jour le jour).

IV. MODELISATION STATISTIQUES

1. Stratégie générale de l'exercice de modélisation

Les propositions faites dans la première partie de ce rapport ont pour objectif d'optimiser le plan d'échantillonnage du réseau météorologique du RNSA de la région Rhône-Alpes. Un pas supplémentaire dans le sens de l'optimisation du fonctionnement du réseau pourrait être réalisé grâce à la mise en œuvre d'une stratégie de modélisation. L'étude présentée dans la

deuxième partie de ce rapport a pour objectif de tester la possibilité de modéliser les données fournies par un capteur sur la base des informations enregistrées au niveau des autres sites. La perspective à long terme serait ainsi de développer une stratégie de collecte de donnée couplant métrologie de constat et modélisation.

L'étude 2 a montré (figure 14) que les propriétés polliniques d'un site donné présentent une grande variabilité interannuelle, mais qu'en revanche la structure du dendrogramme (ou en d'autres termes la distance pollinique qui lie chaque paire de site) est stable d'une année à l'autre. Cela suggère fortement qu'il est plus prometteur de modéliser un site⁶ donné sur la base d'un ou plusieurs autres sites enregistrés la même année, plutôt que sur la base des enregistrements réalisés au niveau de ce site les années précédentes. Différents modèles fondés sur ce constat ont été développés :

- le premier vise à modéliser la réponse d'un site par **interpolation** des concentrations polliniques enregistrées au niveau des sites voisins,
- les autres ont pour objectifs de **modéliser la réponse d'un site sur la base de la réponse d'un autre site par ajustement de l'amplitude, ou de l'amplitude et de la phase** (*i.e.* décalage temporel). Ce dernier type de modélisation implique de dériver la valeur des paramètres d'ajustements sur la base des années précédant la mise en œuvre du modèle. Plusieurs déclinaisons de ces modèles ont été étudiées et leur validité respective a été évaluée.

2. Cadre de l'étude et méthode

2.1. définition des limites de l'étude

L'étude "modélisation" porte sur le même set de données que l'étude 2, c'est à dire **9 sites, 9 taxa** et 3 années (**2003-2005**). Dans le cas des modèles nécessitant un ajustement de paramètres, les années 2003-2004 permettent de dériver la valeur optimale des paramètres, et la modélisation est réalisée pour l'année 2005. Ainsi, il est possible d'évaluer la validité des modèles par comparaison des résultats du modèle avec les valeurs effectivement enregistrées en 2005.

2.2. modélisation par interpolation

Il est possible de modéliser le déroulement temporel de la concentration pollinique d'un site par

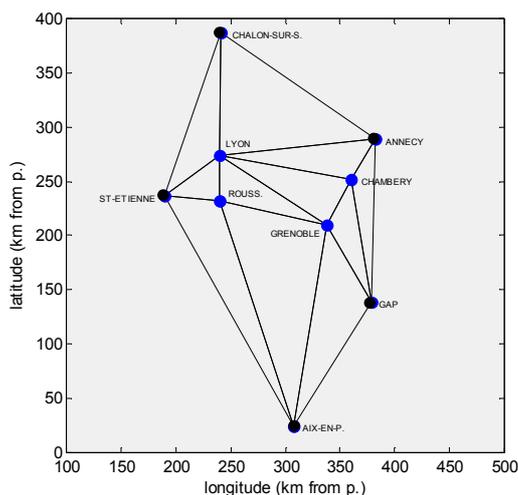


Figure 15. Seuls 4 sites (en bleu) peuvent être modélisés par interpolation

interpolation de la concentration quotidienne enregistrée au niveau des sites voisins. L'interpolation réalisée dans l'étude présentée est basée sur une fonction cubique, mais le programme matlab permet d'implémenter tout type de fonction et ce choix pourra donc être corrigé à l'avenir. Etant donné la structure du réseau, seuls les enregistrements polliniques de quatre sites peuvent être l'objet d'une modélisation par interpolation : Lyon, Roussillon, Chambéry et Grenoble (figure 15).

⁶ Remarque relative au vocabulaire utilisé. Les termes "modéliser un site" ou "modéliser la réponse d'un site" sont souvent utilisés dans ce rapport et signifient : "modéliser les déroulements temporels de concentration pollinique enregistrés au niveau d'un site".

La figure 16 illustre le résultat de la modélisation par interpolation du décours temporel de la concentration de frêne au niveau de Lyon en 2003. Dans ce cas, le DPP-MD (DPP entre la cinétique Modélisées et les Données effectivement enregistrées) est de 92%.

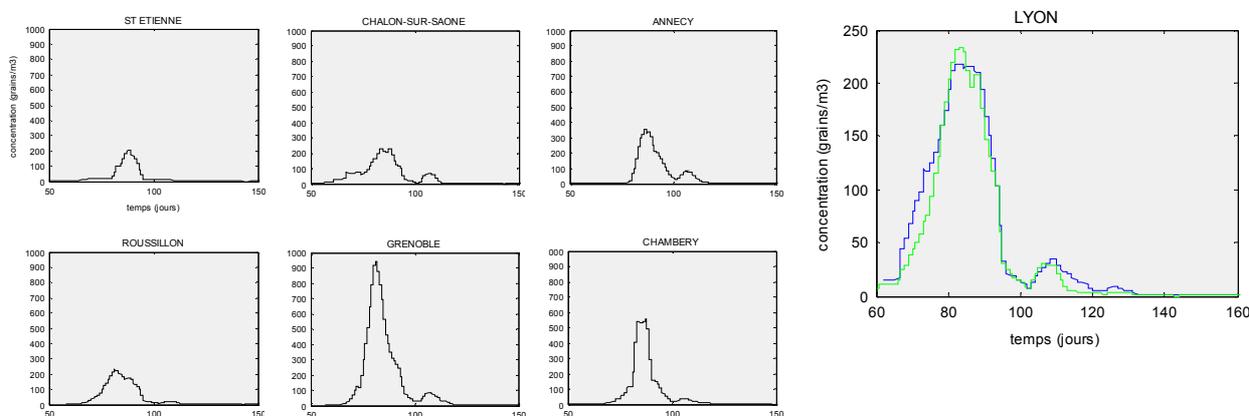


Figure 16. Exemple de modélisation par interpolation du décours temporel de la concentration de frêne au niveau de Lyon en 2003. L'enregistrement modélisé (vert), l'enregistrement original (bleu) ainsi que les enregistrements des sites voisins ayant servi à l'interpolation (noir) sont présentés. La valeur de DPP-MD est de 92%.

2.3. modélisation par ajustement de la phase et/ou de l'amplitude.

Deux types de modèles ont été testés : l'un vise à modéliser la réponse d'un site sur la base de la réponse d'un autre site ("site de référence") par ajustement de l'amplitude (un facteur multiplicatif est appliqué à la valeur quotidienne de concentration) et l'autre par ajustement de l'amplitude et de la phase (la courbe représentant le décours temporel est décalée le long de l'axe du temps). Ces ajustements sont illustrés grâce à deux exemples présentés figure 17.

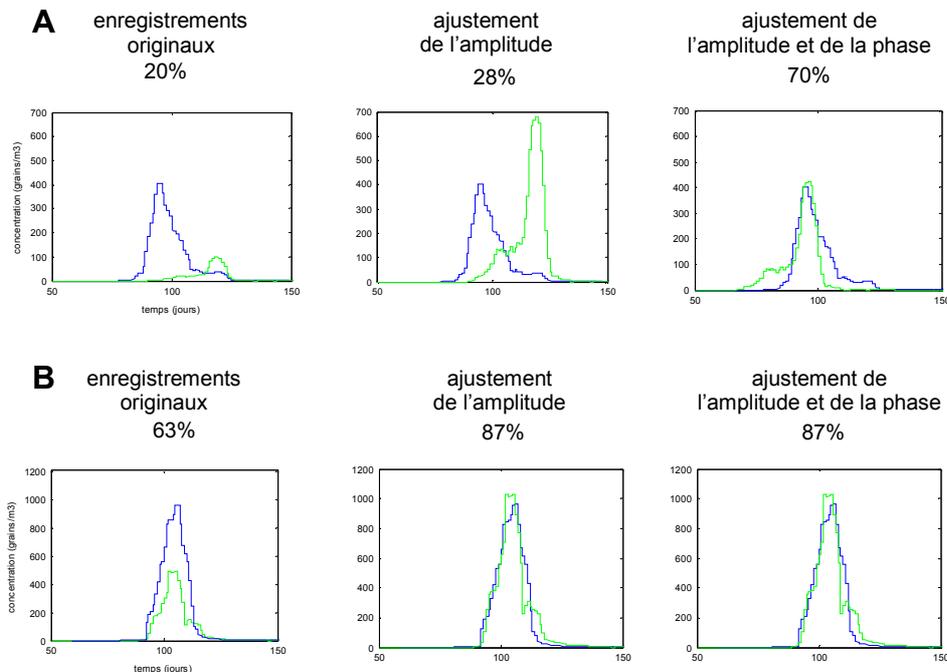


Figure 17. Deux exemples (A et B) d'ajustement de l'amplitude ou de l'amplitude et de la phase. L'exemple A illustre que l'ajustement de l'amplitude seule peut être peu efficace. L'exemple B illustre le cas de deux sites présentant une haute DPP et pour lesquels l'ajustement de la phase ne conduit pas à une augmentation du DPP.

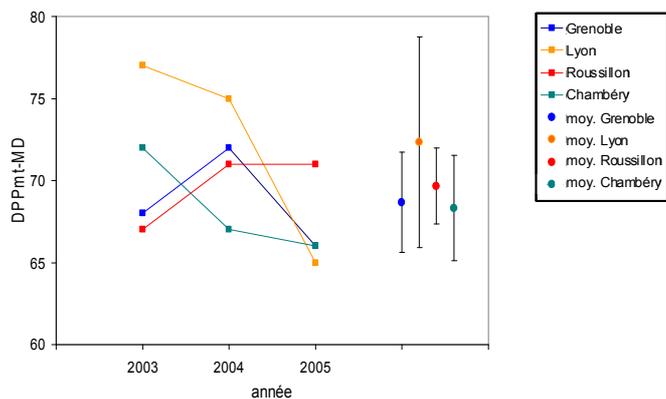
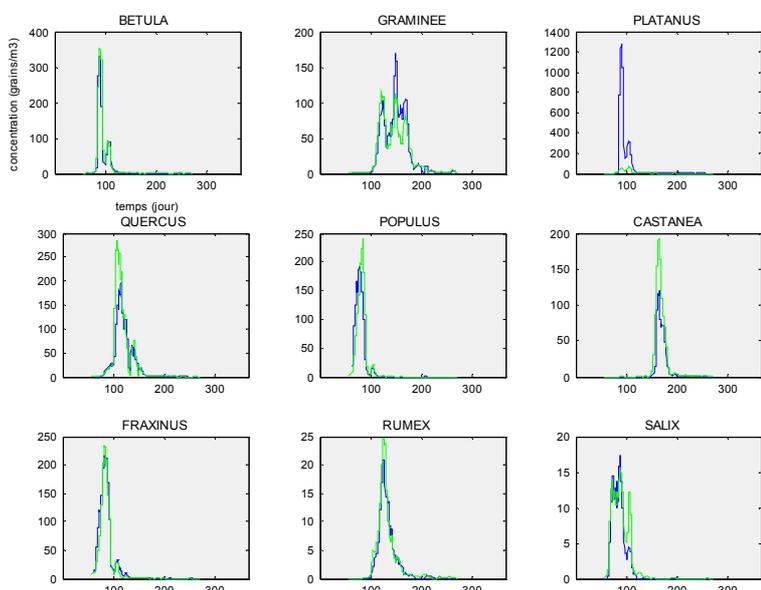
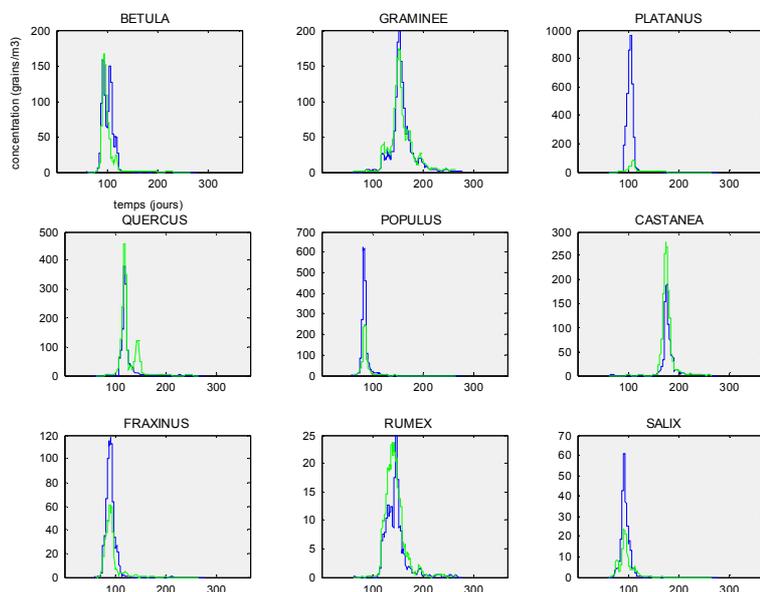


Figure 18. Evolution de la valeur de DPPmt-MD pour les 4 sites pouvant être modélisés par interpolation de 2003 à 2005. Les valeurs de DPPmt-MD sont comprises entre 66% et 77%.

Figure 19. Représentation graphique du résultat de la modélisation par interpolation ayant menée à la plus grande (A) et à la plus petite (B) valeur de DPPmt-MD au cours des années étudiées. Les décours temporels modélisés sont représentés en vert et les décours temporels effectivement enregistrés sont représentés en bleu.



A. Modélisation par interpolation de Lyon en 2003. La valeur de DPPmt-MD est de 77% et les valeurs de DPP-MD de chaque taxon sont : betula 88%, graminee 84%, platanus 16%, quercus 83%, populus 80%, castanea 72%, fraxinus 92%, rumex 88%, salix 86%.



B. Modélisation par interpolation de Lyon en 2005. La valeur de DPPmt-MD est de 66% et les valeurs de DPP-MD de chaque taxon sont : betula 73%, graminee 87%, platanus 15%, quercus 77%, populus 58%, castanea 72%, fraxinus 68%, rumex 80%, salix 60%.

Dans le cadre de l'étude réalisée, la valeur des paramètres d'ajustement est calculée de manière à maximiser le DPP des années 2003 et 2004 entre la réponse modélisée et la réponse effectivement enregistrée. Ce paramètre est noté DPP-MD₂₀₀₃₋₂₀₀₄. Ces paramètres sont ensuite appliqués aux enregistrements du site de référence de l'année 2005. Les résultats de cette modélisation peuvent ensuite être comparés aux enregistrements effectivement réalisés au niveau du site modélisé en 2005 par le calcul de la valeur du DPP-MD₂₀₀₅. La modélisation est réalisée pour chaque taxon indépendamment.

Deux déclinaisons de ces modèles ont été testées. La première consiste à sélectionner le site de référence parmi les sites du réseau. La deuxième consiste à baser la modélisation non pas sur la réponse d'un site effectivement enregistré, mais sur la réponse du site à modéliser, modélisée par interpolation. Ainsi, dans ce cas deux étapes successives sont nécessaires : la réponse du site est modélisée par interpolation, puis cette réponse est l'objet d'un ajustement de phase et/ou d'amplitude.

Cinq modèles ont donc été étudiés, ils sont basés respectivement sur :

- L'interpolation de la réponse des sites voisins (modèle **Int**)
- L'ajustement de l'amplitude sur la base de l'enregistrement modélisé par interpolation (modèle **IntA**)
- L'ajustement de l'amplitude sur la base de l'enregistrement d'un site donnée du réseau (modèle **siteA**)
- L'ajustement de l'amplitude et de la phase sur la base de l'enregistrement modélisé par interpolation (modèle **IntAP**)
- L'ajustement de l'amplitude et de la phase sur la base de l'enregistrement d'un site donné du réseau (modèle **siteAP**)

Dans le cas des quatre derniers modèles, les paramètres d'ajustement sont dérivés sur la base des années 2003-2004 et la modélisation est mise en œuvre sur les données de l'année 2005.

3. Résultats – Discussion

3.1. interpolation

Il a été montré dans la partie III que les propriétés polliniques variaient en relation avec la topographie de la zone et définissaient ainsi des zones homogènes. Dans ce contexte une approche de modélisation envisageable est l'interpolation. Néanmoins, la validité de l'approche dépend de la dimension des zones d'homogénéité et de la densité du plan d'échantillonnage. En effet, l'interpolation au sein d'une zone homogène fournira un modèle plus représentatif des données, qu'une interpolation entre deux sites de zones homogènes différentes.

Modélisation de la réponse d'un site par interpolation, quelle est l'incertitude engendrée ?

De 2003 à 2005, les quatre sites dont les décours temporels sont susceptibles d'être modélisés par interpolation présentent des valeurs de DPP_{mt}-MD comprises entre 66% et 77% (figure 18). Aucun des 4 sites n'est modélisé avec plus de succès que les autres. Sous le postulat que les années 2003 à 2005 sont représentatives des années suivantes, la substitution de la modélisation à la métrologie au niveau d'une de ces 4 stations permettrait d'obtenir des informations avec une fiabilité reflétée par la gamme de DPP_{mt}-MD 66-77%. La variabilité de cette gamme sera d'autant mieux caractérisée qu'elle sera définie sur la base de plus nombreuses années. La figure 19 illustre concrètement la signification de ces valeurs. Elle présente le résultat de la modélisation par interpolation ayant mené à la plus grande (figure 19

A) et à la plus petite (figure 19 B) valeur de DPPmt-MD au cours des années étudiées. Si la cinétique de pollinisation de certains taxa est modélisée avec une DPP-MD supérieure à 90%, il est notable que le platane présente les valeurs de DPP-MD les plus faibles quels que soient l'année et le site considérés. Cet effet peut être lié à une aire de répartition ou des propriétés de transport spécifiques qui entraînent de fortes variations de la concentration locale et rend délicat une simple approche par interpolation. Alternativement, cela peut être lié à l'influence de platanes locaux se trouvant à proximité d'un ou plusieurs capteurs impliqués dans l'interpolation.

Cartographie du décours temporel de concentration basée sur l'interpolation

Il est possible de modéliser par interpolation le décours temporel de concentration pollinique au niveau de chaque point de l'espace inclus dans le polygone délimité par les stations périphériques. La figure 20 illustre le résultat d'une telle approche. Elle représente la concentration d'ambrosie modélisée en chaque point de l'espace à trois dates différentes. Les stations de mesure sont représentées par des croix, l'algorithme d'interpolation utilisé contraint le modèle de telle manière que la concentration représentée au niveau des stations correspond à la concentration effectivement mesurée. La validité des valeurs de concentrations interpolées dépend du plan d'échantillonnage et des dimensions des zones d'homogénéité pollinique. Ainsi, il est vraisemblable que le modèle représente mieux les données au cœur d'une zone homogène qu'entre des zones homogènes pour une densité équivalente de capteur. Par exemple, les valeurs de concentrations modélisées à proximité de Gap sont à considérer avec la plus grande prudence. Sur la base des données disponibles la seule estimation de la validité du modèle Int consiste à substituer la modélisation à la métrologie au niveau d'un site, comme cela a été fait précédemment (Cf figure 18). Mais cette approche reste très partielle et les propriétés du modèle mises en évidence au niveau de ces 4 sites ne peuvent vraisemblablement pas être étendues directement à l'ensemble de la zone d'intérêt. La validité d'une représentation telle que celle illustrée par la figure 20 est donc délicate à estimer et ne peut pas être utilisée en tant que telle. Néanmoins, puisque la concentration représentée au niveau des stations correspond à la concentration effectivement mesurée, ce type de représentation est tout de même intéressante puisqu'elle permet de visualiser (sous la forme d'une animation) la concentration en fonction du temps et de l'espace d'une manière facile à interpréter (pour comparaison Cf. figure 6).

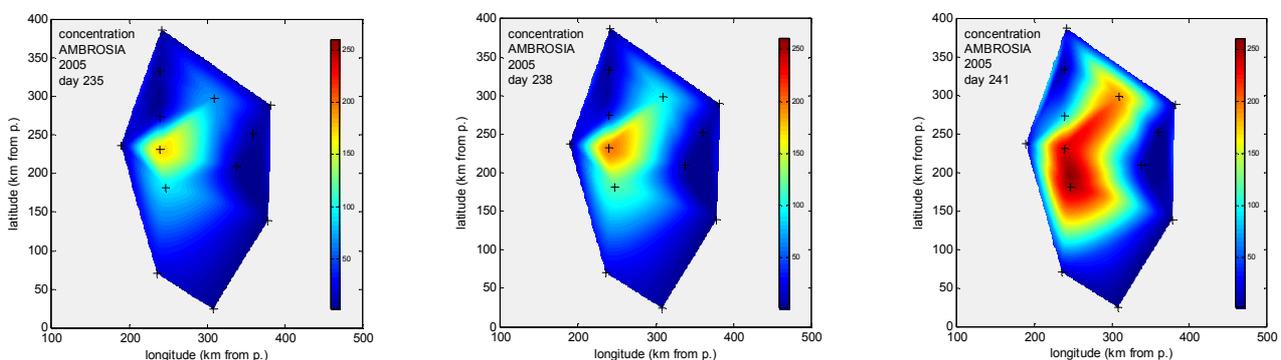


Figure 20. Cartographie de la concentration d'ambrosie modélisée par interpolation sur la base des enregistrements de 13 stations de mesure à 3 dates différentes (235°, 238° et 241° jour de l'année) de l'année 2005. La position des stations de mesure est représentée par une croix noire.

Pour conclure le modèle Int permet de modéliser la réponse d'un site (Lyon, Grenoble, Chambéry ou Roussillon) avec une valeur de DPP_{mt-MD} comprise entre 66% et 77%. Il conviendrait de caractériser plus spécifiquement la signification de ces valeurs avec les utilisateurs des données du réseau, allergologues et aérobiologistes, mais il est vraisemblable que 66% soit une valeur de DPP_{mt-MD} considérée insuffisante pour envisager de substituer la modélisation à la métrologie au niveau d'un site. De plus, cette valeur n'est définie que sur la base de 3 années et l'estimation de sa variabilité est donc à considérer avec prudence. Il est possible qu'une des limitations du modèle Int soit sa faible capacité à rendre compte des spécificités locales, les modèles présentés dans la suite de ce document visent précisément à considérer la spécificité d'un site grâce à l'ajustement de certains paramètres caractéristiques des cinétiques de concentration pollinique : la phase et/ou d'amplitude.

3.2. modèles basés sur l'ajustement de la phase et/ou de l'amplitude

Avant de présenter les résultats des différentes approches de modélisation appliquées aux sites jugés les plus pertinents (section 3.3), les propriétés générales des modèles sont étudiées sur la base de 6 paires de sites sélectionnées au hasard et de 9 taxa. Les objectifs sont essentiellement de comparer les modèles basés sur l'ajustement de l'amplitude seule ou de l'amplitude et de la phase, et d'évaluer la fiabilité de la stratégie qui consiste à ajuster les paramètres sur la base des années 2003 et 2004 afin de les appliquer en 2005.

Effets de l'ajustement de l'amplitude

La figure 21 A illustre l'effet de l'ajustement de l'amplitude. Elle représente la variation de la DPP de chaque paire de site causée par l'ajustement de l'amplitude en fonction de la valeur DPP avant ajustement. Chaque point représente une paire de site et un taxon. Si d'une manière générale l'ajustement de l'amplitude cause une augmentation du DPP d'autant plus élevée que le DPP original est faible, ces deux paramètres montrent une faible corrélation ($R^2=0.1$) et l'ajustement en amplitude n'augmente le DPP de certaines paires de sites que de quelques pourcents. Il semble donc que l'ajustement de l'amplitude seule ne soit pas une méthode qui soit fructueuse pour toutes les paires de sites et tous les taxa.

La figure 21 B représente la DPP_{2005} en fonction de la $DPP_{2003-2004}$ dans le cas des données originales (noir) et dans le cas des données ajustées en amplitude de manière à maximiser la $DPP_{2003-2004}$ (rouge). Chaque point représente une paire de sites et un taxon. En ce qui concerne les données originales (noir), la valeur de DPP_{2005} varie linéairement avec la valeur de $DPP_{2003-2004}$ ($R^2=0.7$, les sites proches polliniquement en 2003 et 2004 le sont toujours en 2005). Cela n'a rien d'étonnant et est à mettre en relation avec la stabilité dans le temps de la structure des clusters polliniquement homogènes évoquée dans l'étude présentée dans la section III. En première approximation, l'effet de l'ajustement de l'amplitude peut être évalué par la comparaison du nuage de points noirs (données originales) et du nuage de points rouges (données ajustées). L'ajustement se traduit par une augmentation du $DPP_{2003-2004}$ (cela est intrinsèque au modèle puisque l'ajustement est réalisé sur la base de ces années) et par une augmentation du DPP_{2005} qui est précisément l'effet désiré. En moyenne l'ajustement de l'amplitude produit une augmentation du $DPP_{2003-2004}$ de 12% et la DPP_{2005} de 9%. Il convient de noter que certaines paires de sites conservent une valeur faible de DPP tant pour la période 2003-2004 que pour 2005, c'est à dire que conformément à ce qui a été mis en évidence précédemment (Cf. figure 21 A) l'ajustement en amplitude peut dans certains cas s'avérer peu fructueux.

Afin de caractériser plus finement l'effet de l'ajustement de l'amplitude, les variations de DPP entre les enregistrements originaux et les données ajustées relatives à 2005 sont représentées

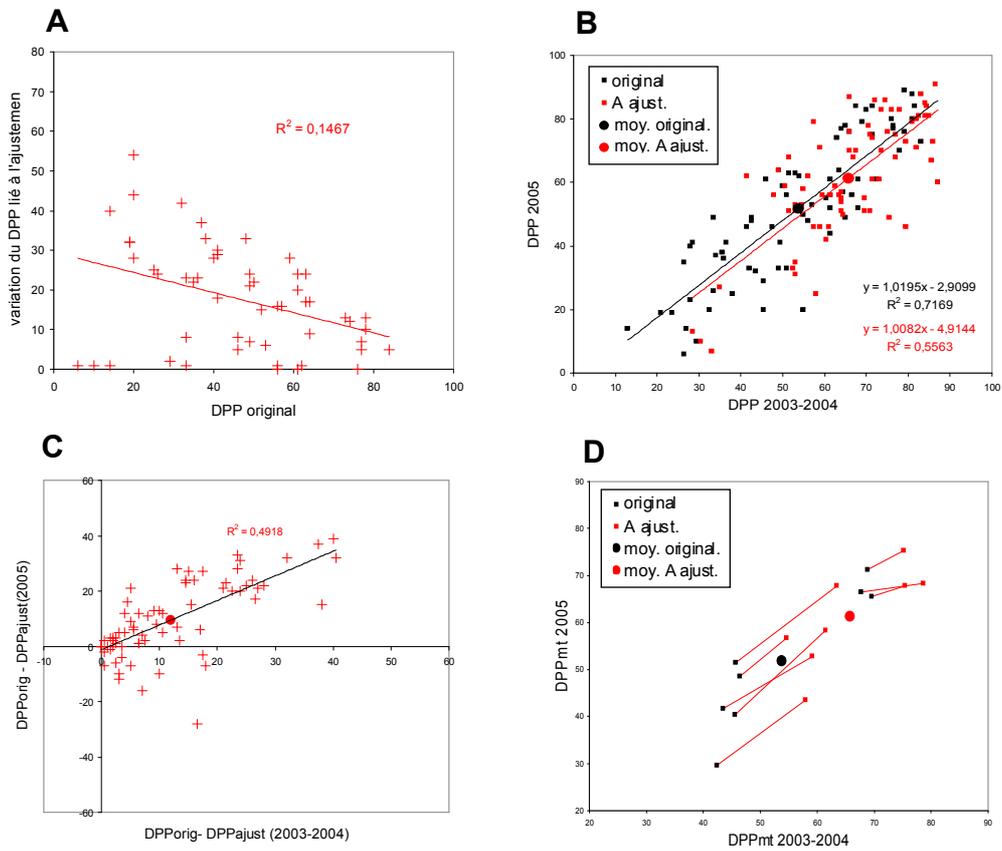


Figure 21 A-D. Effets de l'ajustement de l'amplitude

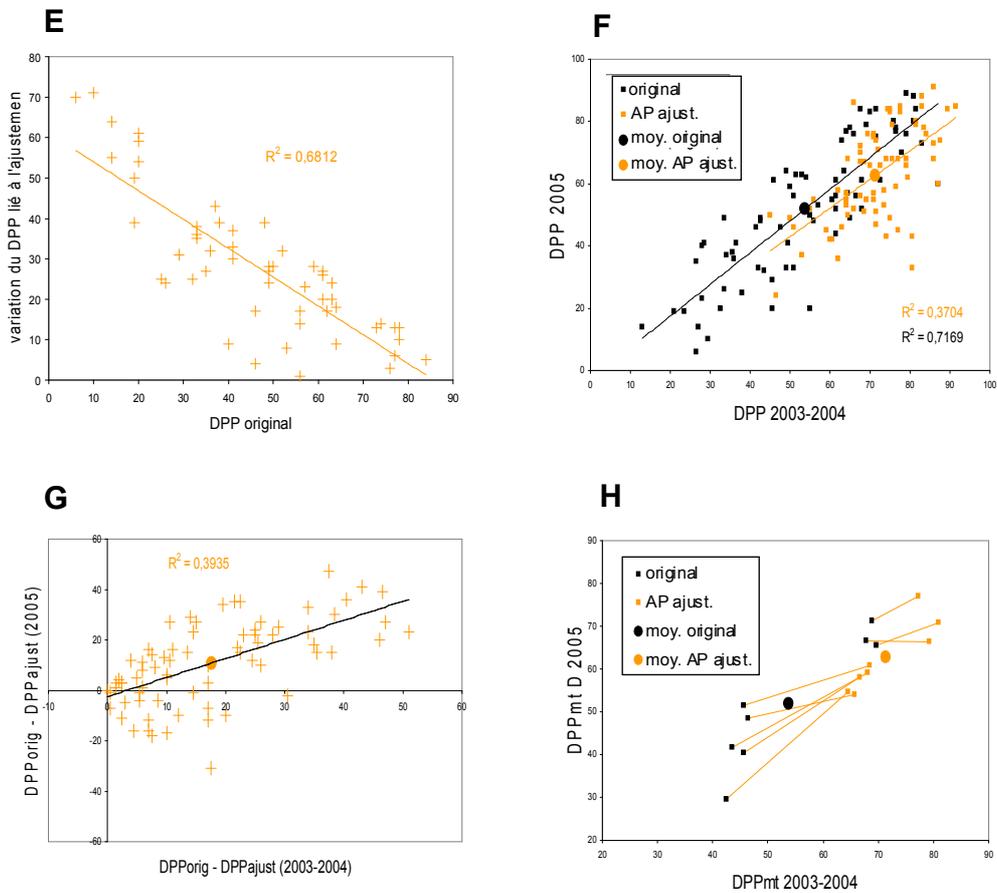


Figure 21 E-H. Effets de l'ajustement de l'amplitude et de la phase

en fonction de celles correspondant à la période 2003-2004 (figure 21 C). Conformément à ce que la figure 21 B illustre, l'ajustement de l'amplitude augmente la $DPP_{2003-2004}$ de tous les taxa de toutes les paires de site considérées, et en général l'augmentation de la $DPP_{2003-2004}$ est associée à une augmentation de la DPP_{2005} . Cela signifie que d'une manière générale la stratégie qui consiste à ajuster l'amplitude sur la base des années 2003-2004 afin de l'appliquer sur l'année 2005 est valable. Ce point est à mettre en relation avec la stabilité des clusters évoquée dans la partie III du rapport, et montre plus spécifiquement que la différence d'amplitude entre les sites est stable d'une année à l'autre. Néanmoins, cela n'est valable qu'en règle générale car dans certains cas (18% pour le set de données considéré) l'application du paramètre ajusté sur les enregistrements de 2005 entraîne une dégradation du DPP_{2005} .

Ainsi, si en général, la stratégie de modélisation qui consiste à dériver un paramètre sur la base de deux ans et l'appliquer l'année d'après est valide, elle peut dans certains cas s'avérer contreproductive. La fiabilité de la méthode serait vraisemblablement augmentée si le paramètre d'ajustement était dérivé sur la base de plus de deux années car cela permettrait de mieux caractériser la relation d'amplitude. De plus il faut tout de même relativiser le "risque" lié à l'utilisation de cette méthode puisque, dans le cas du set de données utilisé, toutes les paires de site présentent une augmentation des DPP_{mt} pour l'année 2005 (figure 21 C).

La stabilité de la relation d'amplitude qui lie deux sites a une implication directe sur la mise en œuvre du modèle. En effet, en pratique, il faut sélectionner le site qui pourra être modélisé l'année suivante avec le plus de succès sur la base des années disponibles. Il convient donc de caractériser la relation entre le DPP_{2005} et le $DPP_{2003-2004}$ des données ajustées afin de définir la stratégie de sélection à adopter (figure 21 A, points rouges). D'une manière générale et comme cela est attendu, le DPP_{2005} est corrélée au $DPP_{2003-2004}$ au niveau des taxa individuels. Ainsi, le choix du site devant être modélisé doit se porter sur le site de la paire fournissant la plus grande valeur de $DPP_{2003-2004}$. Néanmoins, le niveau de corrélation est moyen ($R^2=0.6$), la stratégie de sélection n'est donc valable que d'une manière générale. Cette observation est confirmée au niveau des DPP_{mt} , puisque par exemple le site présentant la $DPP_{mt2003-2004}$ maximale n'est pas celui qui a la DPP_{mt} maximale en 2005 (figure 21 D; points rouges).

Effets de l'ajustement de l'amplitude et de la phase

La variation de DPP de chaque paire de site causée par l'ajustement de la phase et de l'amplitude est corrélée à la valeur de DPP des données avant ajustement (figure 21 E; $R^2=0.7$). Contrairement à l'ajustement de l'amplitude seule (figure 21 A), l'ajustement de l'amplitude et de la phase permet d'augmenter la DPP de toutes les paires de site y compris de celles présentant un DPP original très faible. La figure 17 A illustre un cas où l'ajustement conjoint de la phase et de l'amplitude s'avère plus fructueux que l'ajustement de l'amplitude seule. Néanmoins, si les deux types d'ajustement ne sont pas aussi efficaces en ce qui concerne les paires de sites présentant de basses valeurs de DPP, ils sont d'une efficacité comparable dans le cas des sites présentant de hautes valeurs de DPP. Cela tient essentiellement au fait que les sites qui ont un DPP élevé présentent généralement des cinétiques polliniques en phase, dans ce cas l'ajustement de la phase ne permet pas d'améliorer la DPP. Un tel cas est illustré figure 17 B.

La figure 21 F montre que le calcul des paramètres d'ajustement sur la base des années 2003-2004 permet de réaliser une augmentation globale de la $DPP_{2003-2004}$ (18% en moyenne) ainsi que de la DPP_{2005} (11% en moyenne). L'augmentation du DPP_{2005} moyen est plus grande dans le cas de l'ajustement de l'amplitude et de la phase que dans le cas de l'ajustement de l'amplitude seule. Cela est essentiellement lié à l'augmentation des DPP correspondant à des sites présentant de faibles valeurs de $DPP_{2003-2004}$. En effet, les valeurs maximum de DPP_{2005} sont similaires pour les deux modèles (91% dans les deux cas). Ce point est conforme à ce

qu'illustre la figure 21 E, c'est à dire que pour les sites présentant avant ajustement un DPP élevé, l'ajustement de la phase ne permet pas d'améliorer le DPP.

La figure 21 G décrit plus précisément la variation de DPP causée par l'ajustement. Elle montre que, si d'une manière générale l'augmentation de la $DPP_{2003-2004}$ est corrélée à celle de la DPP_{2005} , l'utilisation des paramètres ajustés sur la base des années 2003-2004 sur les enregistrements de 2005 cause une diminution de la DPP dans 37% des cas et conduit à un coefficient de corrélation faible ($R^2=0.4$). Ce résultat illustre que la relation de phase entre deux sites est peu stable d'une année à l'autre et que la stratégie qui consiste à dériver la différence de phase qui lie deux sites sur la base de seulement deux années pour l'appliquer l'année suivante est risquée. Cette propriété émerge aussi au niveau de DPP_{mt} où une paire de site montre une diminution de la $DPP_{mt_{2005}}$ (figure 21 H).

L'instabilité de la relation de phase entre deux sites au cours du temps est à l'origine de la faible corrélation entre la $DPP_{2003-2004}$ et DPP_{2005} des données ajustées (figure 21 F; nuage de points orange, $R^2=0.4$). En effet, pour une valeur donnée de $DPP_{2003-2004}$, le DPP_{2005} présente une large gamme de valeur, il est donc impossible de présumer sur la base des données 2003-2004 quel site pourra être modélisé avec le plus de succès en 2005.

Pour conclure, l'ajustement de l'amplitude et de la phase fonctionne mieux que l'ajustement de l'amplitude seule sur les paires de sites présentant une valeur de DPP faible. En revanche, les performances des deux modèles sont similaires dans les cas où les DPP sont déjà élevées car d'une manière générale, les sites proches polliniquement sont en phase. Ainsi, puisque l'objectif est d'obtenir des valeurs de DPP maximales, l'ajustement de la phase s'avère d'un intérêt limité. De plus, la relation de phase qui lie une paire de site semble être variable d'une année à l'autre. La valeur de la phase lorsqu'elle est dérivée sur la base des seules années 2003 et 2004 ne permet pas de modéliser de manière robuste l'année 2005. La relation d'amplitude qui lie deux sites est elle aussi variable mais à un niveau qui ne justifie pas de rejeter le modèle. Il faut noter que le calcul des variables d'ajustements n'a été réalisé que sur la base de deux années, l'accumulation de données supplémentaires dans les années à venir pourrait permettre de raffiner la valeur des paramètres d'ajustement.

3.3. Mise en œuvre des différents modèles

Les différentes stratégies de modélisation ont été mises en œuvre afin de comparer les différents modèles de manière concrète. Ainsi, la modélisation des décours temporels de concentration des différents taxa lors de l'année 2005 a été réalisée au niveau des sites de Lyon, Roussillon, Chambéry et Grenoble. Les résultats présentés ne portent que sur ces 4 sites car les modèles nécessitant l'interpolation des valeurs de concentration ne peuvent être mis en pratique qu'au niveau de ces sites. Les modèles qui ne nécessitent pas d'interpolation ont été testés sur d'autres sites mais se sont de toute façon avérés plus performants sur ces 4 sites que sur les autres sites du réseau et les résultats ne sont donc pas présentés.

La figure 22 représente pour chacun des sites et chaque modèle, la valeur de $DPP_{mt-MD_{2005}}$ en fonction de la valeur de $DPP_{mt-MD_{2003-2004}}$.

Le modèle Int basé sur l'interpolation simple des valeurs de concentration de sites voisins ne repose pas sur l'ajustement de paramètre et il est donc attendu que la valeur de $DPP_{mt-MD_{2005}}$ soit sensiblement identique à celle du $DPP_{mt-MD_{2003-2004}}$. Néanmoins, et comme discuté précédemment, la valeur de DPP_{mt-MD} présente une certaine variabilité interannuelle, vraisemblablement liée à une relative instabilité de la structure des clusters polliniquement homogènes.

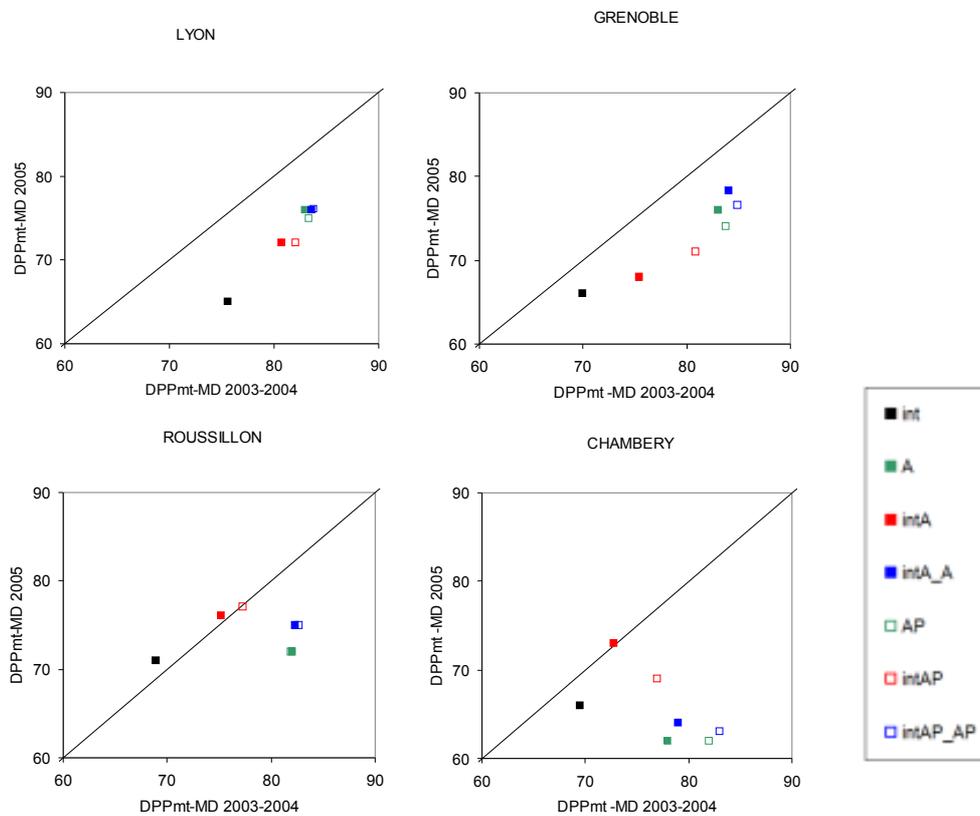


Figure 22. la valeur de $DPPmt-MD_{2005}$ est représentée en fonction de la valeur de $DPPmt-MD_{2003-2004}$ pour chacun des quatre sites et chacun des 7 modèles implémentés

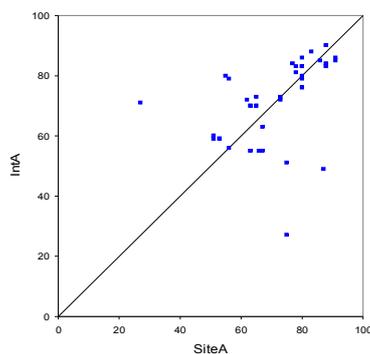


Figure 23. La valeur de $DPP-MD_{2005}$ du modèle IntA est représenté en fonction de celle du modèle SiteA. Cette figure montre que selon le taxon considéré le modèle le plus performant est différent.

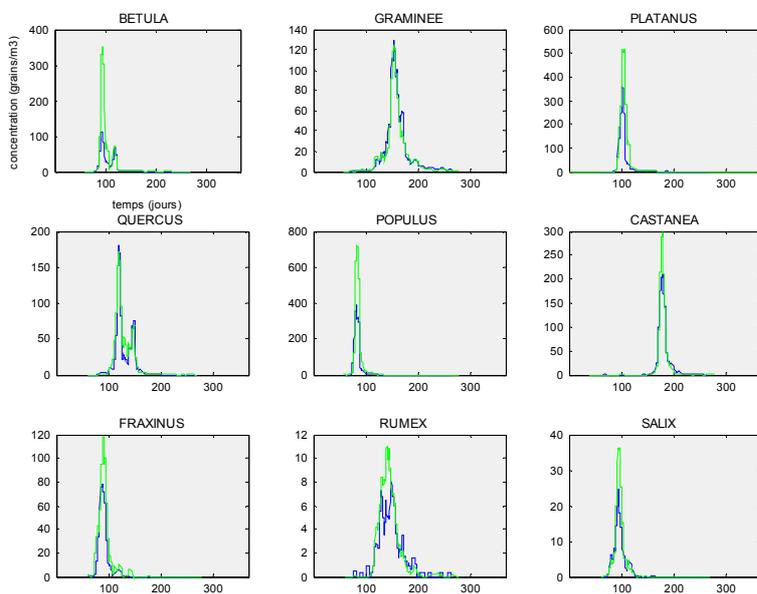


Figure 24. Illustration graphique de la modélisation des décours temporels de tous les taxa réalisée grâce au modèle IntA-SiteA au niveau du site de Grenoble en 2005. La sélection du site et l'ajustement des paramètres du modèle ont été réalisés 'en aveugle' (sans connaître les enregistrements effectifs de 2005). Les décours temporels modélisés sont représentés en vert et les décours temporels effectivement enregistrés sont représentés en bleu. La valeur de $DPPmt-MD$ est de 78%. Les valeurs de $DPP-MD$ de chaque taxon sont de : betula 60%, graminee 91%, platanus 75%, quercus 88%, populus 67%, castanea 88%, fraxinus 75%, rumex 80%, salix 81%.

Le modèle IntA repose sur l'ajustement de l'amplitude du décours temporel modélisé par interpolation sur le décours temporel effectivement enregistré au niveau de ce site. Cet ajustement est réalisé au cours des années 2003 et 2004 et la valeur du paramètre d'ajustement est appliquée au décours temporel de concentration modélisé par interpolation en 2005. Par comparaison avec les résultats du modèle Int, il apparaît que l'ajustement de l'amplitude conduit à une augmentation du DPPmt-MD₂₀₀₃₋₂₀₀₄ et à une augmentation de la DPPmt-MD₂₀₀₅ au niveau des 4 sites. Néanmoins, il convient de noter que le site présentant la plus grande valeur de DPPmt-MD₂₀₀₃₋₂₀₀₄ est Lyon alors que c'est en fait Roussillon qui en 2005 présente la plus grande valeur de DPPmt-MD. Tout cela est conforme aux propriétés de l'ajustement de l'amplitude déjà discutées dans la section 3.2.

Le modèle IntAP est similaire à IntA mais la phase est ajustée en même temps que l'amplitude. Par comparaison avec IntA, il apparaît que l'ajustement de la phase permet d'augmenter le DPPmt-MD₂₀₀₃₋₂₀₀₄ d'autant plus que la DPPmt-MD originale (Int en l'occurrence) est faible. Il est intéressant de remarquer que Roussillon, qui présente un DPPmt-MD₂₀₀₃₋₂₀₀₄ très élevé dans le cas du modèle IntA (carré rouge plein), l'ajustement de la phase ne permet d'augmenter la DPPmt-MD₂₀₀₃₋₂₀₀₄ que de façon marginale (carré rouge ouvert). Cela illustre que l'ajustement de la phase n'a d'intérêt que dans le cas où les DPP-MD sont grandes. De plus, si la DPPmt-MD₂₀₀₃₋₂₀₀₄ est toujours supérieure à celle du modèle IntA, la valeur de DPPmt-MD₂₀₀₅ est supérieure pour 2 sites, égale pour un site et inférieure pour Chambéry. Cela illustre le caractère "risqué" des modèles basés sur l'ajustement de la phase et montre qu'ils doivent être utilisés avec circonspection.

Le modèle SiteA consiste à modéliser le décours temporel de concentration pollinique au niveau d'un site sur la base du décours temporel d'un autre site. Le site de référence est choisi pour chaque taxon indépendamment. Le site sélectionné est le site qui présente la plus haute valeur de DPP₂₀₀₃₋₂₀₀₄. Cette stratégie de sélection s'appuie sur le constat que la valeur de DDP₂₀₀₅ est corrélée à la valeur de DDP₂₀₀₃₋₂₀₀₄. Cette propriété a déjà été discutée et est illustrée par la figure 21 B. La comparaison des résultats correspondant aux modèles IntA et SiteA, montre qu'alors que le modèle SiteA permet toujours d'obtenir les valeurs les plus élevées de DPPmt-MD₂₀₀₃₋₂₀₀₄, les valeurs de DPPmt-MD₂₀₀₅ ne sont supérieures à celles de IntA qu'au niveau de 2 sites sur les 4 étudiés : Lyon et Grenoble. Le set de données étudié est restreint et ne permet pas de conclure quant aux propriétés respectives de IntA et SiteA, néanmoins, IntA semble être un modèle un peu plus robuste que SiteA. Il est possible de faire l'hypothèse que, parce que la réponse modélisée par interpolation est le reflet de la réponse de plusieurs sites, la relation d'amplitude entre les concentrations polliniques enregistrées et celles modélisé par interpolation est plus stable d'une année à l'autre que la relation d'amplitude entre le site à modéliser et un seul autre site.

Le modèle SiteAP est construit sur la même base que SiteA mais, la phase est ajustée conjointement à l'amplitude. L'effet spécifique de l'ajustement de la phase dans le cas de SiteAP est similaire à celui de IntAP et ne sera donc pas discuté à nouveau.

Modèles composites : IntA-SiteA et IntAP-SiteAP. La figure 23 représente la valeur de DPP-MD₂₀₀₅ du modèle IntA en fonction de celle du modèle SiteA. Elle montre que selon les taxa, un modèle peut-être meilleur que l'autre (*i.e.* tous les points ne se trouvent pas sur la droite $y=x$). Il est donc envisageable de modéliser la réponse d'un site en utilisant selon les taxa le modèle IntA ou SiteA. Pour chaque taxon, le choix du modèle est fait sur la base des valeurs de DPP-MD₂₀₀₃₋₂₀₀₄. Ce modèle composite est baptisé IntA-SiteA (carrés pleins bleus). Un autre modèle basé sur IntAP et SiteAP est aussi testé. Les propriétés de ces modèles sont les mêmes que celles des modèles leur servant de base, elles ne seront donc pas discutées davantage.

Modélisation des décours temporels de pollen en 2005, quelle incertitude ?

Lorsque toute la procédure de modélisation et de sélection des sites est réalisée sans avoir connaissance des données de 2005, **IntA** permet de modéliser les enregistrements de Lyon en produisant une valeur de DPP_{mt}-MD de 72% et des valeurs de DPP-MD pour les taxa individuels compris entre de 49% et 85%. Il est notable que dans le cas du modèle IntA, c'est le site de Roussillon qui présente la plus grande valeur de DP_{mt}-MD en 2005, mais sur la base des années 2003-2004 c'est le site de Lyon qui est sélectionné (Figure 22). Le modèle **SiteA** permet de modéliser les enregistrements de Lyon en produisant une valeur de DPP_{mt}-MD de 76% (le DPP-MD est compris entre 56% et 91%). Enfin lorsque le modèle **IntA_SiteA** est implémenté, le site sélectionné est Grenoble, la valeur de DPP_{mt}-MD est de 78% et les valeurs de DPP-MD appartiennent à la gamme 60%-88%. La figure 24 est une illustration graphique de la modélisation des décours temporels de tous les taxa réalisée grâce à ce modèle, elle permet d'appréhender qualitativement la valeur de DPP-MD.

4. Conclusions

Différents modèles ont été mis en œuvre. Le modèle Int permet de modéliser la réponse d'un site (Lyon, Grenoble, Chambéry ou Roussillon) avec une valeur de DPP_{mt}-MD comprise entre 66% et 77%. Si ce modèle présente l'intérêt de pouvoir modéliser chaque point de l'espace, il faut noter qu'il est délicat d'évaluer la validité d'une telle approche sur la base des données disponibles. Les modèles qui consistent à dériver le décours temporel de concentration pollinique à partir d'un autre site en ajustant l'amplitude grâce à un facteur dérivé sur la base des enregistrements réalisés les années précédentes se sont montrés assez efficaces. Les modèles IntA, SitesA et IntA_SiteA ont permis de modéliser les cinétiques de concentration des différents taxa au niveau d'un site en produisant des valeurs de DPP_{mt}-MD comprises entre 72% et 78%. En l'état actuel du set de données, il est difficile d'estimer lequel de ces modèles est le plus performant. Il est vraisemblable que les modèles se montreront plus robustes lorsque le paramètre d'ajustement (l'amplitude) pourra être dérivé sur la base de plus de deux années. Enfin, les modèles basés sur l'ajustement de l'amplitude et de la phase sont à utiliser avec prudence car il semble que la relation de phase qui lie deux sites ne soit pas stable d'une année à l'autre.

La modélisation statistique est une approche qui semble envisageable à moyen terme et dont la mise en œuvre est susceptible d'optimiser la collecte des données en terme économique (en limitant le nombre de capteurs) et de représentativité géographique (mise en œuvre de cartographie). Il faut noter que si l'option de la modélisation était retenue, la stratégie de développement du réseau métrologique devrait être adaptée en conséquence. En effet, un site d'implantation de capteur devrait alors être choisi en fonction de sa capacité à servir de site de référence pour la modélisation d'autres sites, plutôt qu'en fonction de sa stricte représentativité métrologique. En outre, puisque la relation entre les sites est stable d'une année à l'autre, il est envisageable d'utiliser des stations temporaires – implantées pour une durée permettant de définir les paramètres d'ajustement propres au site d'implantation – afin de multiplier le nombre de sites modélisés et de caractériser l'incertitude propre à chacun de ces sites.

Les modèles exposés demandent à être raffinés avant d'être inclus dans une stratégie de collecte de données couplant métrologie de constat et modélisation. Le développement des modèles ne repose pas uniquement sur des compétences statistiques mais nécessite une forte implication des allergologues qui sont de fait les premiers utilisateurs des données polliniques. En particulier, il est essentiel que soit défini le niveau d'incertitude qui peut être toléré sur les décours temporels modélisés.

V. CONCLUSION GENERALE

Les études présentées avaient pour objectif d'évaluer la possibilité d'optimiser le fonctionnement du réseau métrologique du RNSA. Deux voies ont été explorées : l'optimisation du plan d'échantillonnage et la modélisation.

L'approche statistique mise en oeuvre a permis de définir des zones polliniquement homogènes au sein de la région Rhône-Alpes. La distribution de ces zones homogènes a servi de base à l'exploration des possibilités d'optimisation du plan d'échantillonnage du réseau métrologique de RNSA. L'étendue des zones homogènes n'a pas pu être caractérisée à cause de la dimension réduite de la zone étudiée. L'étude pourrait donc être avantageusement étendue à une région plus vaste. Dans le cadre de l'optimisation du réseau métrologique, la possibilité d'utiliser à la fois des capteurs de Hirst et des intercepteurs de Cour afin de bénéficier de leurs avantages complémentaires, pourrait être étudiée avec attention.

Une autre voie d'optimisation du fonctionnement du réseau qui a été explorée est la modélisation. L'objectif était d'estimer dans quelle mesure il est possible de dériver les données au niveau d'un site à partir des enregistrements réalisés au niveau d'un autre site. L'idée étant, à terme, de coupler métrologie de constat et modélisation. Les différents modèles implémentés devront être raffinés au cours des années à venir, mais l'étude montre tout de même que la modélisation purement statistique est une approche envisageable. Bien que ce soit en dehors du champ de ce rapport, il est notable que la modélisation déterministe de concentration de pollen constitue une autre voie prometteuse à long terme.

Les études réalisées mettent en évidence certains aspects du réseau grâce à une approche purement statistique, il va de soi qu'ultimement le choix de la stratégie de surveillance aérobiologique est conditionné par les besoins sanitaires, la répartition de la population et les contraintes économiques. Il convient en particulier de définir d'une manière collégiale qu'elle est l'incertitude qui peut être admise sur les décours temporels modélisés.

L'ensemble des études présentées a été réalisé grâce à un programme développé sous matlab pour les besoins de l'étude. Cet outil informatique pourra être utilisé dans le futur pour actualiser les paramètres des modèles, pour étudier une autre zone géographique ou d'autres taxa (voire des données d'une autre nature). De plus, c'est un outil flexible et les choix statistiques qui pourraient être l'objet de discussions (fonction utilisée pour l'interpolation des données, le mode de regroupement utilisé pour l'analyse en cluster...) pourront être facilement ajustés.