Conservatoire et Jardin botaniques Genève

Certificat complémentaire en géomatique 2017-2018
Rapport de Stage

UNIVERSITÉ DE GENÈVE

Travail de Certificat

Liste Rouge des trachéophytes du canton de Genève : vers une automatisation de la génération des statuts de menace

Frédéric Alexandre Sandoz

Sous la direction des

Dr. Pascal Martin Dr. Anthony Lehmann

Table des matières

1.Préambule **4**
2. Introduction **4**
2.1 Contexte 4
2.2 Objectifs 4
2.3 Buts d'une liste rouge 5
2.4 Réactualisation de la liste rouge des trachéophytes du canton de Genève ; quelques concepts importants 6
2.4.1 Critère A2c : Réduction de la taille de la population 6
2.4.2 Critère B2 : Répartition géographique 6
3. Méthodologie **7**
3.1 Outils informatiques 7
3.2 Préparation des données de base avec ArcGIS 8
3.3 Préparation des données floristiques 10
3.4 Préparation des différents types de mailles 12
3.5 Préparation des différentes cartes de regroupement de milieux naturelles 14
3.6 Préparation des données pour l'attribution d'un statut de menace provisoire selon le critère A2c 16
3.7 Attribution des statuts de menaces provisoires selon le critère A2c 20
3.8 Critère B2. Répartition géographique et créations des localités 21
3.8.1 Génération des localités 21
3.8.2 Définition d'une localité 21
3.8.3 *ModelBuilder* pour la création des prélocalités 22
3.9 Création semi-automatisée des localités 32
4. Résultats obtenus via *ModelBuilder* **34**
4.1 Exemples de résultats de créations de prélocalités obtenus avec les 4 méthodes 34
4.2 Exemple concret de création de localités avec la méthode N°4 37
5. Problèmes rencontrés avec *ModelBuilder* : quelques recommandations **38**
6. Passage sur FME **39**
6.1 Attribution des statuts de menace provisoires selon le critère A avec FME 39
6.2 Préparation des données sur FME 41
7. Discussion **44**
8. Conclusion **46**
9. Bibliographie **46**
7. Annexe **47**

1. Préambule

L'intérêt de ce travail est double. Outre l'acquisition du Certificat complémentaire en géomatique et le fait de pouvoir exercer à l'échelle 1:1 différents outils acquis tout au long de ce cursus, le but était aussi de participer, sous la forme d'un stage, à l'élaboration de la nouvelle Liste Rouge des trachéophytes du Canton de Genève. Dès lors, ce travail est d'avantage un guide méthodologique, expression quelque peu prétentieuse, ancré dans les besoins de cette entreprise difficile qu'est la conception d'une liste rouge qu'un rendu strictement académique. Le fait que ce stage soit d'avantage lié à la pratique a pour conséquence un rendu différant quelque peu du schéma conventionnel normalement attendu dans ce type d'exercice. Ainsi, les lecteurs et correcteurs de ce travail s'apercevront assez vite que, souvent, méthodologie et résultats ne forment qu'un continuum.

2. Introduction

2.1 Contexte

Dans le canton de Genève, ce sont les Conservatoire et Jardin Botaniques de la Ville de Genève (CJB) qui font office d'autorité compétente pour la définition des statuts de protection cantonaux des espèces végétales. En 2006, une première liste rouge a été publiée **(Lambelet-Haueter et al., 2006)**. Puis, dans le cadre du programme « *Monitoring de la flore et des milieux du canton de Genève* » en partenariat avec la Direction Général de l'Agriculture et de la Nature (DGAN), les CJB se sont vus confier l'actualisation de la Liste Rouge des trachéophytes du canton avec une échéance fixée à 2019. Les CJB doivent ainsi réactualiser les statuts de menace de près de 1250 taxons avant cette date. Dans cette optique, les CJB souhaitent mettre en place une procédure d'attribution des statuts de menace qui soit le plus automatique possible. Cela est d'autant plus crucial, que le jeu de données à traiter est conséquent ; pas loin d'un demi-million d'entrées pour l'ensemble du « Grand Genève » !

2.2 Objectifs

A la base, le but de ce stage était de mettre en place un système expert d'attribution des statuts de menaces à l'aide du logiciel eCognition. L'objectif final était d'arriver à une procédure qui puisse traiter un maximum de taxons tout en permettant de mettre en exergue les espèces pour lesquelles les données seraient insuffisantes ou dont le statut aurait changé depuis la première Liste Rouge. Malheureusement, la constitution des données de base n'était pas achevée au commencement du stage. Pis, les *Model Builder* construits précédemment dans le but de faciliter la création des premières étapes de traitements de données ont révélé quelques écueils qu'il fallait corriger pour pouvoir avancer. Dès lors, les objectifs du stage ont été modifiés.

Les « nouveaux » objectifs sont les suivants :

- Amélioration et correction des modèles existants dans le but de trier les entrées des 1000 premiers taxons retenus pour la réactualisation de la Liste Rouge.
- Attribution d'un statut de menace provisoire sur ces taxons de façon automatisée pour le critère A2c.
- Création de nouveaux modèles permettant l'automatisation de la génération de localités pour l'ensemble des taxons retenus nécessaire à l'attribution du critère B2.
- Transition et adaptation des différents modèles obtenus via les *ModelBuilder* sur FME.

Pour y parvenir, l'excellent travail d'**Alison Lacroix (2016)** portant sur l'adaptation des critères UICN pour le canton de Genève et le Certificat de géomatique de **Matthieu Chevalier (2016)** ont servi de socle à ce travail. Bien que ce document repose en partie sur les bases établies par Matthieu Chevalier en 2016, il s'en distingue néanmoins par un approfondissement des méthodes permettant la génération des localités.

2.3 Buts d'une liste rouge

Malgré la petitesse de son territoire, le canton de Genève dispose d'une mosaïque de milieux fort diversifiés permettant l'établissement d'une diversité végétale foisonnante qui fait le bonheur des botanistes locaux **(Theurillat et al., 2011)**. En 1832 déjà, G.F. Reuter s'exprimait en ces termes à propos de l'extraordinaire diversité floristique de canton : « *Il est peu d'endroits aussi intéressants pour le botaniste que les environs de Genève. Il en est peu, en effet, où la végétation soit aussi variée. Aux portes mêmes de la ville, dans les plus courtes promenades, on trouve un grand nombre d'espèces dont plusieurs ne laissent pas d'être très rares.* » Mais près de deux siècles plus tard, le constat est amer. Force est de constater que « notre » flore n'est pas épargnée par le recul de la biodiversité observé au niveau mondial **(Bornand et al., 2016)**. En Suisse, 95 % des prairies et pâturages maigres ont ainsi disparu depuis le début du XXe siècle suite principalement à l'intensification de l'agriculture **(Fischer et al., 2015)** pour être remplacés par des prairies grasses composées d'espèces banales et ubiquistes **(Gossner et al., 2016 ; Bornand et Sager, 2017)**. Alors que le rythme d'extinction des espèces semble s'être emballé **(Ceballos et al., 2017)**, la conception de listes rouges n'a probablement jamais été aussi importante. A la façon d'un secouriste arrivant sur les lieux d'un grave accident, une liste rouge oriente, donne l'ordre des priorités quant aux secours à apporter. Elle donne un état de l'urgence de la situation. Mais surtout, elle nous aide à prendre conscience de l'érosion de la biodiversité végétale souvent lente et pernicieuse. En effet, les stations et les individus qui les composent n'ont de cesse de diminuer petit à petit et de façon insidieuse. Il y a alors une sorte d'habituation à voir bon nombre d'espèce être de moins en moins fréquentes **(Bornand et Sager, 2017)**. Froidement, les listes rouges enregistrent ce lent déclin et devraient nous faire prendre conscience de cette lente mais implacable dégringolade de notre patrimoine naturel.

En Suisse, les listes rouges représentent un instrument légal pour la désignation de biotopes dignes de protection. La nouvelle liste rouge des plantes vasculaires de Suisse **(Bornand et al., 2016)** invoque, entre autres, les raisons suivantes quant à l'utilité d'un tel document :

- Base pour la définition d'espèces prioritaires nationales ;
- Critère d'évaluation de la qualité des biotopes ;
- Désignation d'espèces cibles dans la planification sectorielle de la protection de la nature ;
- Développement des stratégies de conservation de la biodiversité en Suisse ;
- Evaluation des mesures d'aménagement du territoire sur l'environnement ;
- Sensibilisation du public à la protection des espèces animales ou végétales.

Bien évidemment, ces différentes considérations peuvent aussi être transposées à l'échelle cantonale.

2.4 Réactualisation de la liste rouge des trachéophytes du canton de Genève ; quelques concepts importants

La création de cette nouvelle liste rouge se base sur la méthodologie établie par l'UICN (Union Internationale pour la Conservation de la Nature) initialement développée pour la protection des espèces animales mais adaptée ici à la taille réduite du canton de Genève et à la protection de son patrimoine botanique **(Lacroix, 2016)**. Cinq critères (A, B, C, D, E) sont combinés pour pouvoir attribuer une catégorie de menace **(Tableau 1)** à chacune des espèces et ainsi déterminer leur probabilité d'extinction.

Tableau 1 : Présentation des catégories de menace de l'UICN adaptées à l'échelle régionale.

Taxons menacés
RE : régionalement éteint
CR : en danger critique d'extinction
EN : en danger
VU : vulnérable
Taxons peu ou pas menacés
NT : quasi menacé
LC/R ou LC : non menacé
Taxons non évalués
DD : données insuffisantes
NA : non applicable
NE : non évalué

Ces critères sont par la suite subdivisés de façon plus précise. Pour chacun des taxons retenus, ces critères prennent en compte la taille des populations et la fluctuation de leur effectif dans un laps de temps déterminé, leur structure spatiale (fragmentation) ou bien encore les modifications de leur aire de distribution. Par soucis de simplification, je ne détaille ici que les deux sous-critères A2c et B2 qui ont motivé ce travail de géomatique.

2.4.1 Critère A2c : Réduction de la taille de la population

L'UICN recommande de mesurer l'évolution de la taille d'une population sur un laps de temps de 10 ans ou 3 générations **(UICN, 2011)**. Pour la réactualisation de la liste rouge, il a été choisi d'estimer l'évolution de la taille des populations de chacun des taxons retenus via leur fréquence d'apparition au sein d'un maillage kilométrique entre les intervalles de temps [1990-2005] et [2006-2019]. Selon le critère A2c régionalisé, autrement dit basé sur une zone d'occupation calculée via une maille kilométrique, une réduction ≥ 80 % correspond à un statut CR, une réduction ≥ 50 % à un statut EN et une réduction ≥ 30 % à un statut VU. Le critère A2 s'applique aux réductions de populations constatées, estimées, déduites ou supposées, dans le passé, lorsque les causes de la réduction n'ont peut-être pas cessées OU ne sont peut-être pas comprises OU ne sont peut-être pas réversibles.

2.4.2 Critère B2 : Répartition géographique

Ce critère concerne la zone d'occupation effective d'un taxon. Régionalisé à l'échelle du territoire genevois, les statuts attribués via ce second critère sont les suivants : un taxon occupant moins de 5 km^2 **et** présent dans 1 localité aura un statut CR, EN s'il est présent dans moins de 10 km^2 **et** aux

maximum 5 localités et VU s'il occupe moins de 15 km^2 **et** au maximum 10 localités. Le calcul de ce critère n'a été effectué que pour les taxons dont la répartition était égale ou inférieure à 15 km^2 sur la base du second intervalle de temps et du maillage de 500 m (voir plus bas). La conception des localités occupera l'essentiel de ce travail de géomatique. De ce fait, de plus amples explications seront données plus loin dans le texte quant à leur création et leur définition.

3. Méthodologie

Comme expliqué en préambule, ce travail ne suit pas scrupuleusement les normes établies pour ce type d'exercices. Ainsi, certains modèles sont d'avantage des résultats en eux-mêmes plutôt que des outils utilisés pour les produire. Cela est particulièrement vrai pour les tests effectués pour la génération des localités. Cependant, quelques sorties de modèles sont intégrées à leur description en guise de résultats préliminaires.

3.1 Outils informatiques

Le principal outil utilisé lors de ce stage est **ArcGIS 10.2.1** développé par ESRI (*Environmental Systems Research Institute*) et plus particulièrement son langage de programmation visuel, ***ModelBuilder***, qui permet l'élaboration de séquences de géotraitements complexes. Dans *ModelBuilder,* un modèle est représenté par un diagramme dans lequel s'enchaînent des séquences de traitement et des d'outils de géotraitements. La sortie d'un traitement est alors utilisée comme entrée d'un autre traitement (**http://pro.arcgis.com/fr/**).

Le deuxième outil utilisé est **FME Workbench 2015.1** est un outil d'extraction, de transformation et de chargement des données géographiques vectorielles et image. L'interface est relativement similaire aux *ModelBuilder* d'ESRI (**Figure 1**).

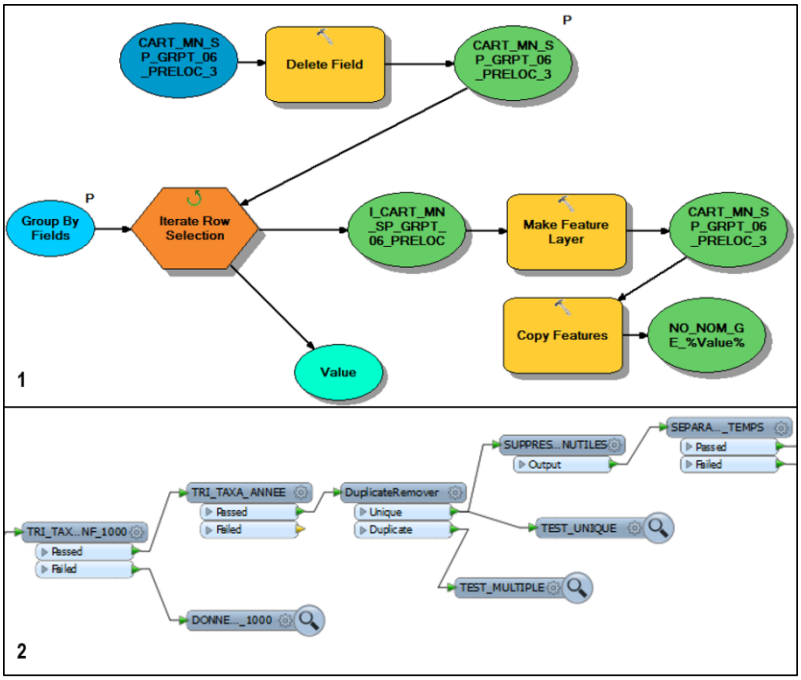


Figure 1 : Comparaison de deux modèles crées sur ArcGIS (1) et FME (2).

3.2 Préparation des données de base avec ArcGIS

Cette section a pour but de décrire les différentes données introduites au sein des *ModelBuilder* ainsi que les modifications nécessaires effectuées sur celles-ci en vue des différents traitements géomatiques.

Deux tables sont à la base de l'ensemble des opérations effectuées sur les données. La première table, **CJB.SIPV_FS_OBS**, contient l'ensemble des notes floristiques transmises par Info Flora et le CBNA (Conservatoire Botanique National Alpin) pour le territoire du « Grand Genève ». Cette table, mise à jour régulièrement, est disponible sur les serveurs de la Ville et de l'Etat de Genève. Elle est composée de nombreux champs mais seul 11 d'entre eux contiennent des informations pertinentes pour la suite des opérations de sélections. Une brève description de ces champs est donnée dans le **Tableau 2** ci-dessous. La table contient aussi de nombreuses autres notes « floristiques » autres que celles liées strictement aux trachéophytes identifiables par le champ CO_GROUPES. Ainsi, sur les 483 569 données que compte la table au 1e décembre 2017, 126 concernent des champignons (**FU**), 2217 des Characées (**CH**), 8488 des bryophytes (anthocérothes (**AN**), hépatiques (**HE**) et mousses (**MU**)), et 17 058 sont des lichens (**LI**). Quelques données (215) ne sont rattachées à aucun des dix « regroupements taxonomiques » et 564 n'apportent aucune information (**XX**). A la date du 1erdécembre 2017, 455 116 données concernent des trachéophytes dont 7653 sont des ptéridophytes (**PT**), 6816 des gymnospermes (**GY**), 104 138 des monocotylédones (**AM**) et 336 509 sont des dicotylédones (**AD**).

Les précisions de ces notes (XY_PRECISION) sont aussi disparates : 1 mètre pour les plus précises à plus de 16 kilomètres pour les plus imprécises d'entre elles.

L'intervalle de temps durant lequel les données ont été collectées est également impressionnant. La plus vielle donnée floristique contenue dans la table **CJB.SIPV_FS_OBS** date de 1612. Il s'agit d'une observation de *Minuartia capillacea* (All.) Graebn. qu'aurait faite Jean BAUHIN, un naturaliste et botaniste bâlois d'origine française, sur Le Reculet. Quant aux données récentes, elles ne cessent d'affluer. Ainsi, au matin du 8 décembre 2017, la table contient 483 784 données soit 215 de plus après la mise à jour hebdomadaire transmise aux CJB par Info Flora.

La deuxième table, **SIPV_FS_SP_LR_19_ATTRIBUTS_TAXONS_GE_PROV_2017_11_15**, contient une liste des taxons qui ont été ou qui sont toujours présents dans le canton de Genève. Comme son nom le laisse à penser, il s'agit d'une table provisoire contenant les taxons qui seront retenus pour la nouvelle liste rouge. Ce travail est effectué par Florian Mombrial, responsable de l'actualisation de la LR. Pour les différents tests, la table utilisée date du 15 novembre 2017 et contient 1013 taxons. Cependant, un certain nombre de ces taxons ne seront pas pris en compte lors de l'élaboration de la liste rouge 2019 et doivent être mis de côté pour la suite des analyses. Sur la base du champ TAXON_RETENU, 11 espèces sont ainsi retirées (*Acer opalus* Mill., *Anagallis foemina* Mill, *Ballota nigra* L. s.str., *Betula pubescens* Ehrh., *Bromus hordeaceus* L. s.str., *Cardamine pratensis* L. [s.str. prov.*], Dianthus superbus* L. s.str., *Eleocharis mamillata* H. Lindb., *Epilobium obscurum* Schreb., *Festuca rubra* L. s.str., *Hieracium glaucinum* Jord.) parce que leur « présence naturelle » sur le territoire genevois a été jugée douteuse. Deux autres taxons (*Cirsium canum* (L.) All., *Castanea sativa* Mill.) ont été retirés sur la base du champ RETENU_LR_19 contenant les espèces pour lesquelles nous souhaiterions une réactualisation de statuts au sein de la nouvelle liste rouge car l'indigénat reste à confirmer pour la première espèce et la seconde est avant tout une espèce cultivée sur le canton. Le

tableau est encore amputé de deux autres taxons *Equisetum* x *meridionale* (Milde) Chiov. et *Minuartia hybrida* (Vill.) Schischk. s.str. qui ne possèdent pas de numéro d'identifiant Info Flora parce qu'ils ne n'ont pas encore été reconnus par Info Flora. La table finale, **TAX_PROV_998_NET**, ne contient dès lors plus que 998 données.

Tableau 2 : Champs devant être conservés pour la poursuite des traitements géomatiques.

NOM DU CHAMP	DESCRIPTION	VALEURS
NO_NOM_IF	Identifiant numérique Info Flora du taxon considéré.	
CO_GROUPE	Regroupement taxonomique auquel a été rattaché le taxon.	• AD : Dicotylédones • AM : Monocotylédones • AN : Anthocérothes • CH : Characées • FU : Champignons • GY : Gymnospermes • HE : Hépatiques • LI : Lichens • MU : Mousses • PT : Ptéridophytes • XX : Aucun regroupement
NO_COMPLET	Nom latin du taxon avec le nom du ou des auteurs.	
DETERMINAVIT_CF	Indication quant à l'incertitude de la détermination.	• N : Détermination certaine de l'espèce • OD : Détermination incertaine de l'espèce (observateur) • OF : Détermination avérée fausse (observateur) • CD : Détermination incertaine de l'espèce (expert) • CF : Détermination avérée fausse (expert) • X : Indéfini / inconnu
PRESENCE	Indication sur la présence de l'espèce.	• + : Présent • N : Non confirmé / non retrouvé • P : Non confirmé, présence vraisemblable • I : Non confirmé, présence incertaine / improbable • A : Station inaccessible • - : Disparu / détruit
DOUTEUX	Indique la confiance avec laquelle le taxon a été identifié.	• 0 : Notes validées • 1 : Notes pas encore validées • 2 : Notes douteuses • 3 : Notes fausses
INTRODUIT	Indication de l'origine de la population de la population.	• N : Origine naturelle / naturalisation • I : Population d'origine incertaine • A : Adventice / Sporadique • S : Echappé / Subspontané • C : Cultivé • R : Réintroduction officielle (origine connue) • O : Introduction sauvage (origine inconnue) • X : Origine de la population inconnue
A_NOTE	Année de la prise de la donnée.	
CO_CANTON	Indication sur le canton ou le département français d'origine de la donnée.	• 01 : Ain • 74 : Haute-Savoie • GE : Genève • VD : Vaud
X et Y	Coordonnées métriques des observations selon le système de coordonnées projetées CH1903+_LV95.	
XY_PRECISION	Ce champ renseigne sur la précision, en mètre, avec laquelle une note a été enregistrée.	

Les champs jugés inutiles de la table **TAX_PROV_998_NET** sont ensuite éliminés. Il ne reste à la fin de l'opération qu'un tableau composé de 998 taxons et de cinq colonnes (OBJECTID, NO_NOM_GE qui est l'équivalent de l'identifiant Info Flora mais pour le canton de Genève, NOM_COMPLET, NO_NOM_IF, RETENU_LR_19). Par ailleurs, la suppression des champs NOM_COMPLET et RETENU_LR_19 n'aurait posé aucun problème pour les traitements géomatiques futurs, mais le fait de connaître l'identité des espèces par leur nom latin permet de repérer plus facilement d'éventuelles erreurs que le seul identifiant NO_NOM_GE, pour peu que l'on connaisse l'écologie et la répartition des taxons dans le bassin genevois. Le champ RETENUS_LR_19 a pour seul but la vérification de la sélection par attribut précédemment effectuée.

3.3 Préparation des données floristiques

La table **TAX_PROV_998_NET** ainsi obtenue est prête « à être croisée » avec la table **CJB.SIPV_FS_OBS** afin d'en extraire les données nécessaires à la création de la nouvelle liste rouge. Le *ModelBuilder* présenté ci-dessous peut être utilisé pour l'extraction des données de la table **CJB.SIPV_FS_OBS.** Il se décompose en trois parties brièvement décrites ci-dessous.

La première (**Figure 2**) étape consiste à sélectionner l'ensemble des trachéophytes contenu dans la table **CJB.SIPV_FS_OBS** au moyen de la sélection par attribut suivante : CO_GROUPE in ('AD', 'AM' , 'GY', 'PT'). Une copie de la sélection est ensuite effectuée. Il est bien sûr tout à fait possible de passer directement à la seconde étape mais copier l'entité permet à la fois de contrôler le résultat de la sélection par attribut mais aussi de stabiliser le *ModelBuilder*. Copier régulièrement les sorties peut aussi faciliter d'éventuelles corrections du *ModelBuilder* en partant directement d'une copie sans avoir à refaire l'entièreté des processus de sélections qui s'avèrent parfois être peu stables.

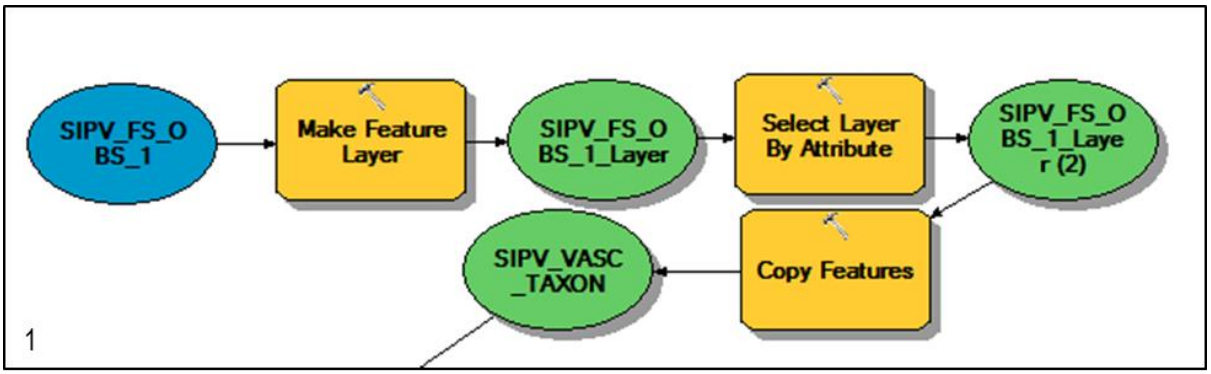


Figure 2 : Sélection des trachéophytes à parti de la table CJB.SIPV_FS_OBS.

La deuxième étape consiste à croiser (*Intersecter* en *franglais*) la sélection avec le maillage **MAILLES_LIMITROPHES_LAC_1000** qui agit ici comme une sorte de tamis à l'aide de l'outil *Intersect*. Les données se trouvant hors du maillage sont ainsi exclues des analyses futures. Plus de détails concernant la création et le type de mailles seront donnés ci-dessous. Les données sont ensuite extraites de la table **CJB.SIPV_FS_OBS** en lui joignant la table **TAX_PROV_998_NET** avec comme champ d'entrée NO_NOM_IF et comme champ de sortie NO_NOM_IF. Cette opération permet aussi d'assigner les numéros d'identifiant genevois renseignant le champ NO_NOM_GE. Cette étape peut facilement prendre une quinzaine de minutes sur une machine dédiée aux traitements géomatiques.

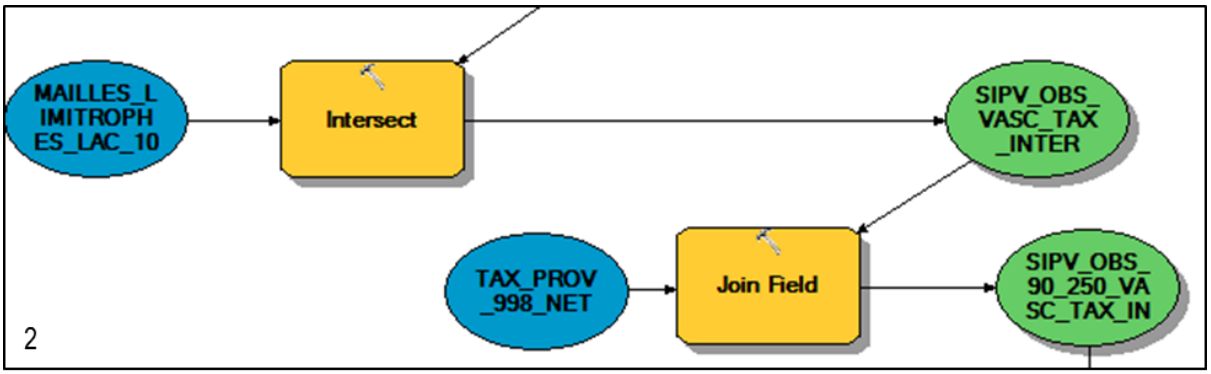


Figure 3 : Intersection des taxons présélectionnés dans l'étape 1 avec le maillage kilométrique.

La troisième partie (**Figure 4**) consiste à trier les données contenues dans la table **CJB.SIPV_FS_OBS** par une série de sélections par attribut. Ainsi toutes les données dont la précision est supérieure à 1000 m sont éliminées (XY_PRECISION <= 1000), de même que les notes prises avant 1990 (A_NOTE > 1989), lors de la première sélection. La deuxième sélection par attribut regroupe un ensemble de quatre sélections qui parachèvent le tri des données floristiques :

DOUTEUX in (0 ,1) AND… sélectionne uniquement les notes validées ou en cours de validation.

PRESENCE in ('+') AND… permet de sélectionner uniquement les observations positives, par opposition aux données négatives qui correspondraient aux stations détruites ou disparues.

DETERMINAVIT_CF in ('-','X') AND… toutes les déterminations fausses ou incertaines sont éliminées par cette sélection.

INTRODUIT in ('A','I','N','O','R','S','X') permet d'éliminer les espèces cultivées.

Une copie est effectuée pour chacune des sélections par attribut et le résultat final est à la fois disponible sous forme de classe d'entité ainsi que sous la forme d'un tableau. La table final **SIPV_1000_90_GE** peut encore être simplifiée en supprimant à l'aide de l'outil *Delete Field* les champs qui ne sont pas nécessaires à la suite des traitements géomatiques mais qui alourdissent le temps de calcul. La table finale devrait ainsi contenir les champs suivants : NO_NOTE, NO_NOM_IF, NOM_COMPLET, A_NOTE, X, Y, XY_PRECISION et NO_NOM_GE.

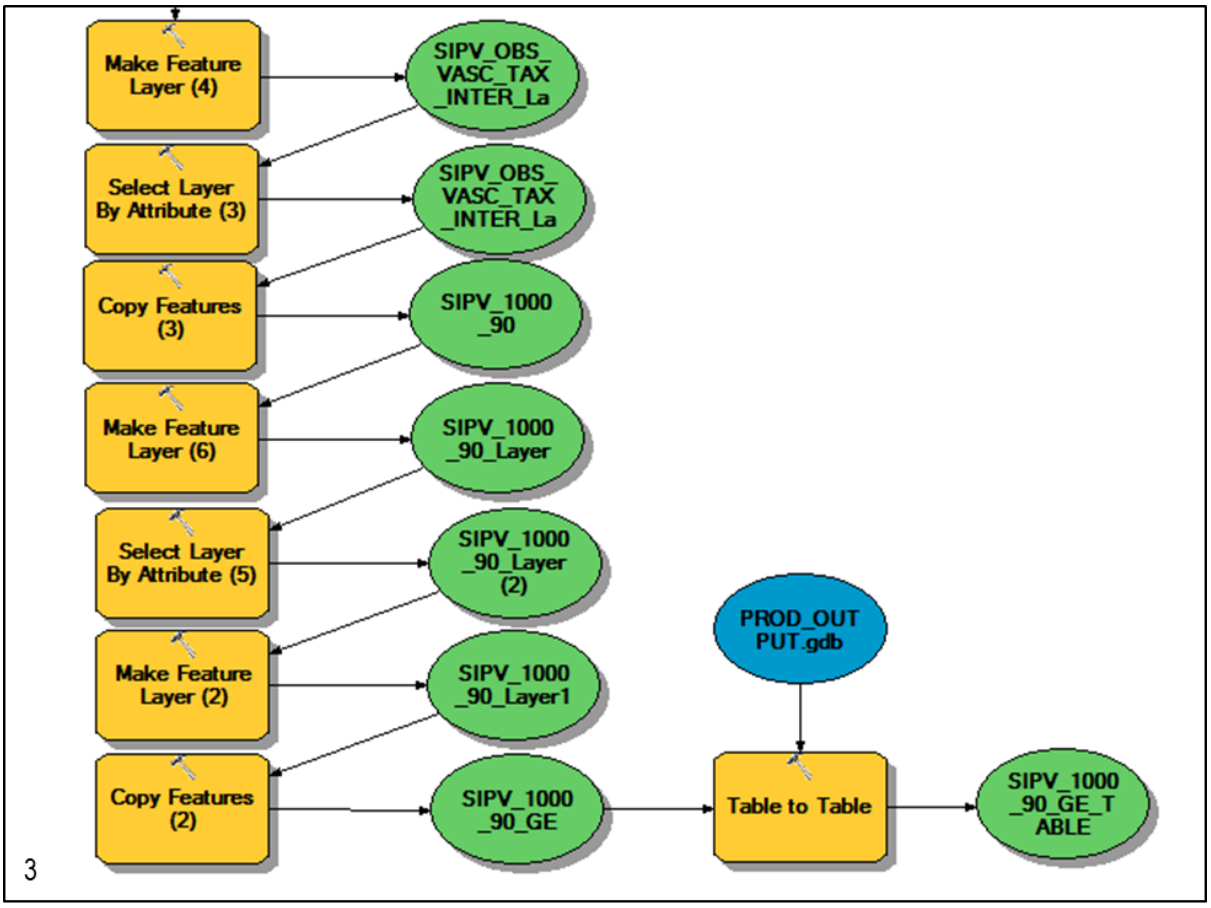


Figure 4 : « Nettoyage » des données floristiques par une suite de sélections par attribut.

3.4 Préparation des différents types de mailles

Deux maillages (**Figure 5**) sont créés pour l'ensemble des traitements géomatiques : **MAILLES_LIMITROPHES_LAC_1000** et **MAILLES_LIMITROPHES_LAC_500_GE**. Comme énoncé ci-dessus, les mailles servent avant tout de tamis pour sélectionner les données de la dition et exclure celles se trouvant en dehors des frontières cantonales et nationales. Elles jouent aussi un rôle central dans les calculs des critères A2c et B2.

Le maillage **MAILLES_LIMITROPHES_LAC_1000** est composé de 357 mailles de 1 km^2. Certaines des mailles débordent largement des frontières cantonale et nationale, ce qui a pour conséquence la prise en compte d'un certain nombre de notes floristiques hors du territoire genevois qui ne seront pas éliminées lors de l'intersection avec la table **CJB.SIPV_FS_OBS**. Cela peut paraître surprenant pour une liste rouge dont la dition est le canton de Genève, mais n'oublions pas que certaines notes avaient une précision kilométrique. Dès lors, une note qui se trouverait sur la frontière et qui aurait une précision de l'ordre du kilomètre pourrait très bien se situer 500 mètres en dehors des frontières cantonales. De plus, les plantes font fi de nos frontières administratives et certaines stations peuvent s'étendre en dehors des limites cantonales. Dans le même ordre d'idée, une espèce peut être considérée comme rare sur le territoire cantonal ce qui ne signifie pas que cela soit également vrai pour l'ensemble du bassin genevois. Dès lors, étendre quelque peu le territoire considéré permet de relativiser la rareté de certaines espèces qui pourraient être rares dans le canton mais d'avantage fréquentes juste en dehors de celui-ci. Outre la préparation des données floristiques, le maillage kilométrique servira aussi à l'évaluation des taxons selon le critère A2c.

Le second maillage, **MAILLES_LIMITROPHES_LAC_500_GE**, est composé de 1282 mailles de 0.25 km^2 construites sur la base du maillage de 1 km^2. Il ne s'agit pas du quadruple des nombres de mailles de 1 km^2, car seul les mailles contenues entièrement ou à l'intersection des frontières cantonale et nationale ont été gardées. Dès lors, le nombre de notes retenues lors de l'intersection de la table **CJB.SIPV_FS_OBS** avec le maillage de 0.25 km^2 est sensiblement moins élevé pour un pas de temps identique. Ce maillage servira à l'évaluation des taxons selon le critère B2.

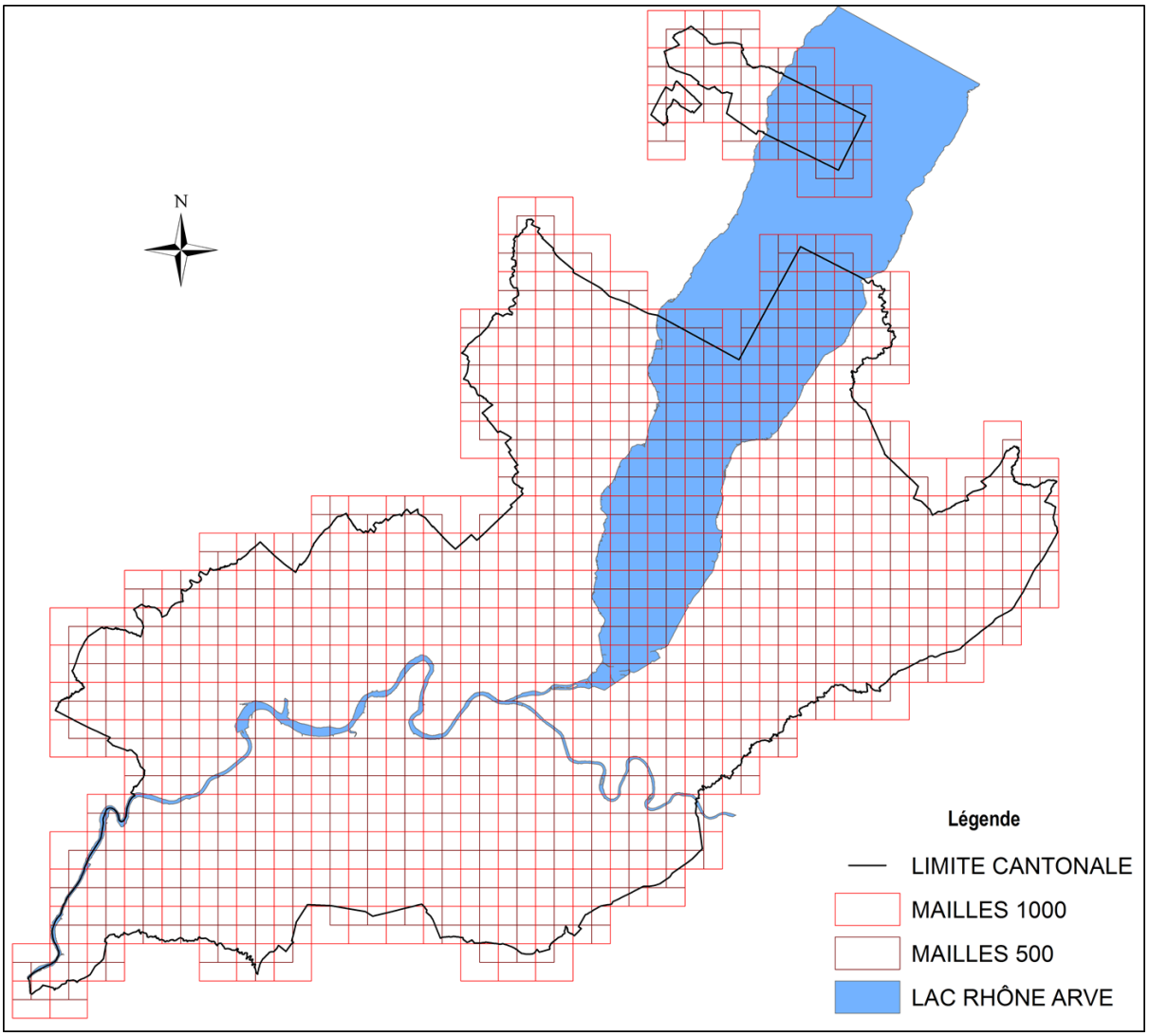


Figure 5 : Emprise des mailles kilométriques et de 500 m sur le territoire genevois.

Pour les deux types de maillages, l'ensemble des mailles lacustres ont été prises en compte. En effet, rien n'empêche la présence d'une *Lemna trisulca* L. qui flotterait en plein milieu du Petit Lac et dont l'observation serait transmise à Info Flora. Ainsi, les mailles doivent couvrir l'ensemble du territoire du canton de Genève dans l'optique de maximiser le nombre de données prises en compte.

3.5 Préparation des différentes cartes de regroupement de milieux naturels

Onze cartes de regroupement de milieux naturels **(Annexe 1)** ont été créées à partir de la classe d'entité **CJB.SIPV_MN_CARTO_BRUTE**. La carte des milieux naturels du canton de Genève sur lesquels se basent les regroupements est mise à jour tous les ans par les CJB et comporte 84 catégories de milieux plus ou moins naturels à fortement anthropiques. En grande partie créée à partir de modélisations, elle continue de faire l'objet de validation et de précisions à partir d'observations effectuées directement sur le terrain.

L'exemple ci-dessous (**Figure 6**) reprend la création du groupement N°1, abrégé **GRPT_01**, qui regroupe 18 milieux naturels liés ou affilés à la végétation aquatique (**Annexe 1**). Une première sélection par attribut est opérée sur la table **CJB.SIPV_MN_CARTO_BRUTE** avec les CODE_MN des 18 milieux naturels regroupés sous le terme de végétation aquatique. Une première carte provisoire, **GRPMT_01selection**, est obtenue avec l'ensemble des milieux naturels sélectionnés. La carte reste cependant un ensemble de milieux naturels indépendants. L'outil *Dissolve* et son option *Create multipart features* décochée permettent d'obtenir la carte du groupement de milieux naturels finale. L'opération sera répétée pour les dix autres groupements manuellement. Un exemple du rendu final est donné par la **Figure 7** ci-dessous.

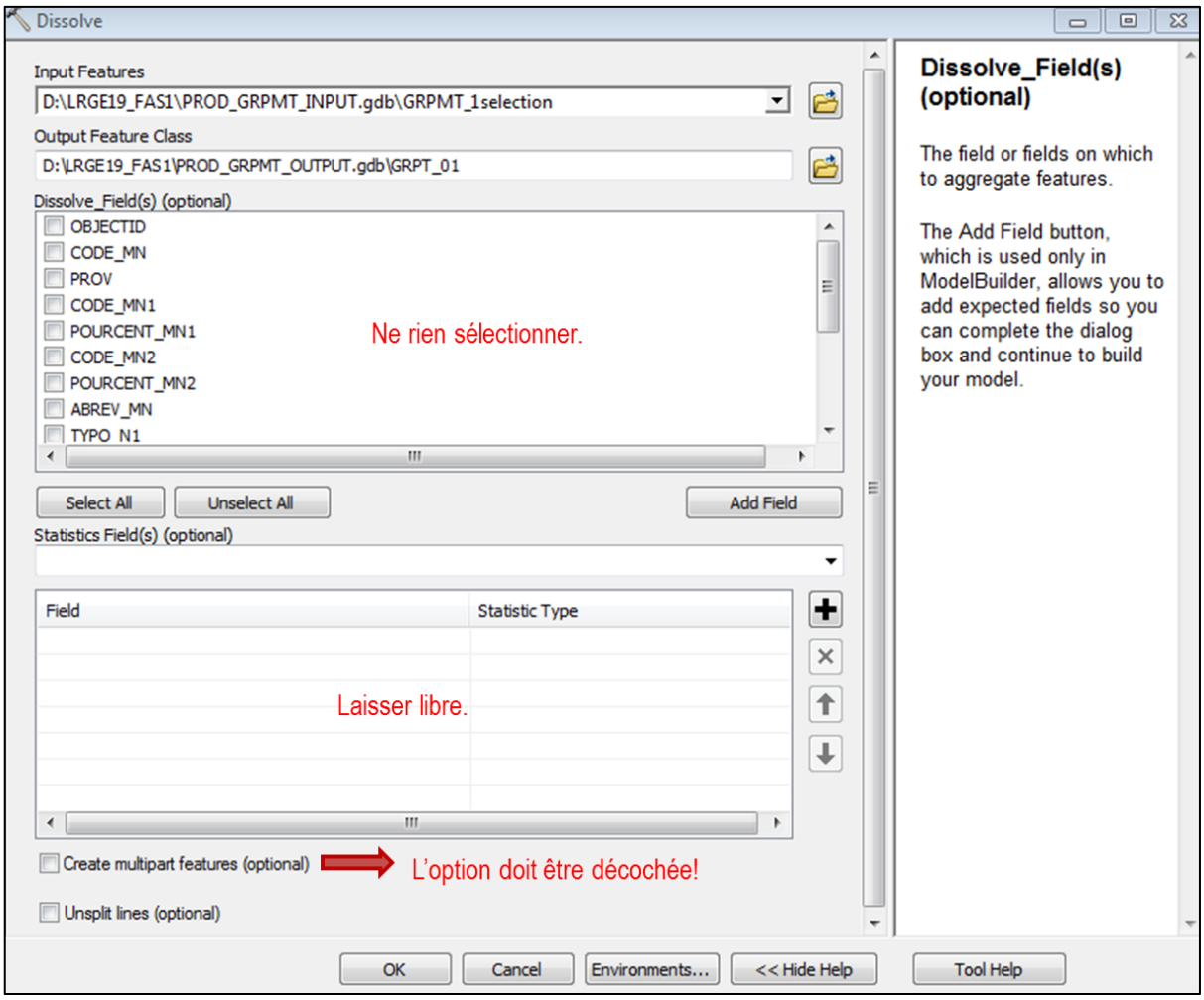


Figure 6 : Création de la carte de groupement de milieux naturels 01à l'aide de l'outil *Dissolve*.

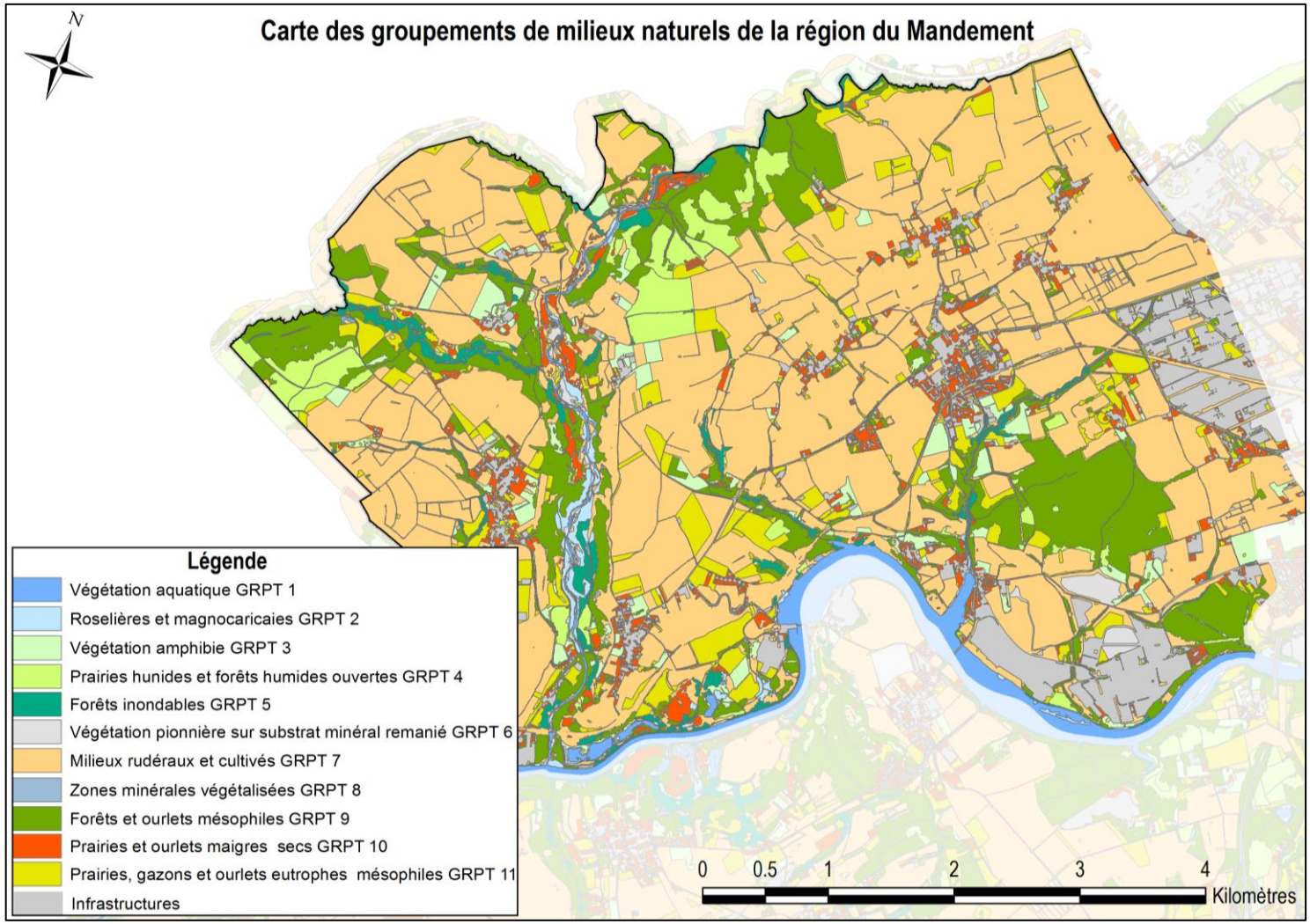


Figure 7 : Carte des groupements de milieux naturels de la région du Mandement. L'ensemble des 11 groupements est représenté.

Un milieu naturel peut faire partie de plusieurs groupements. Par exemple, la vigne appartient au groupement 7 (milieux rudéraux et cultivés) mais aussi aux groupements 10 et 11 (prairies et ourlets maigres/secs ; prairies, gazons et ourlets eutrophes/mésophiles). Certains groupements comprenant des milieux naturels rares ont été complétés avec des milieux naturels plus communs ou écologiquement proches. Par exemple, le groupement 10 (prairies et ourlets maigres /secs) regroupe, entre autres, les pinèdes et les pelouses sèches (*Xerobromion*), milieux rares à l'échelle cantonale, mais aussi les pelouses arborées (prairies, gazons et ourlets eutrophes/mésophiles), les vergers à hautes tiges (prairies, gazons et ourlets eutrophes/mésophiles) ou bien encore la végétation des sols argileux calcaires (milieux rudéraux et cultivés).

3.6 Préparation des données pour l'attribution d'un statut de menace provisoire selon le critère A2c

Le critère A2c peut être calculé à partir de la table **SIPV 1000_90_GE** précédemment obtenue via le *ModelBuilder* N°2.

La première étape (**Figure 8**) consiste à séparer les notes floristiques en deux intervalles de temps distincts. Le premier couvre un laps de temps de 16 ans [1990–2005] et le second une période de 12 ans [2006–2017] à la date du traitement. Pour s'assurer que la première dichotomie a été correctement effectuée, une copie de chacune des sélections est effectuée. Cela permet aussi de garder une certaine stabilité du *ModelBuilder*. En effet, lorsque le modèle est lancé plusieurs fois, il arrive fréquemment que la séparation des notes floristiques en deux intervalles de temps ne s'effectue pas correctement. Le fait d'utiliser l'outil *Make Feature Layer* suivi directement par une copie de la classe d'entité obtenue par la sélection par attribut permet de pallier, en partie, à ce problème.

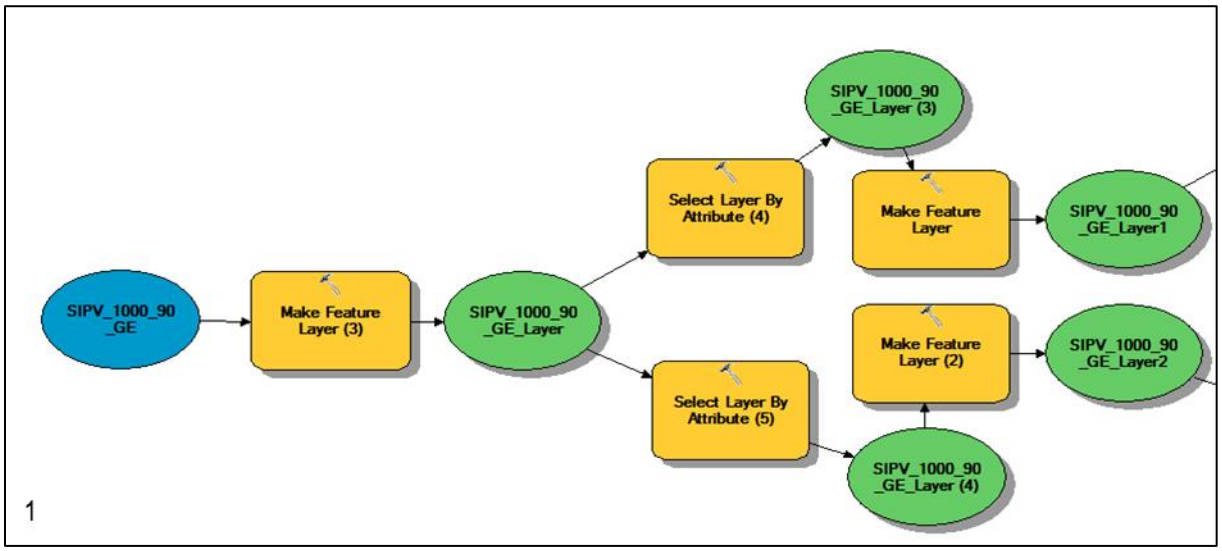


Figure 8 : Séparation des données floristiques en deux intervalles de temps distincts.

La deuxième étape du *ModelBuilder* N°2 peut être décomposée en trois processus. Le premier processus (2a), effectué à l'aide de l'outil *Tabulate Intersection*, permet d'attribuer le numéro d'identifiant d'une maille 1000 (COORGE_ID) aux taxons se trouvant à l'intérieur de son périmètre. Le résultat de cette opération est visible dans l'exemple présenté ci-dessous (**Figure 9**) où le numéro d'identifiant genevois (NO_NOM_GE) d'*Aceras anthropophorum* (L.) W.T.Aiton est rattaché à

l'identifiant (COORGE_ID) de la maille 1000 dans laquelle se trouve le taxon. L'outil *Tabulate Intersection* permet aussi d'éliminer d'éventuels doublons au sein d'une même maille.

Table

TABINT_1000_06_17_1000

OBJECTID *	COORGE_ID	NO_NOM_GE	PNT_COUNT	PERCENTAGE
4905	42	11	1	0.729927
4997	43	11	1	0.729927
5106	44	11	4	2.919708
5562	50	11	1	0.729927
6775	65	11	1	0.729927

Figure 9 : Réattribution du numéro d'identifiant genevois NO_NOM_GE.

Le second processus (2b) n'est effectué que sur les données floristiques du deuxième intervalle de temps [2006-2017]. L'outil *Tabulate Intersection* fait intervenir le maillage **MAILLES_LIMITROPHES_LAC_500_GE** qui agit à nouveau à la manière d'un tamis. Finalement, le troisième processus (2c) permet d'obtenir la fréquence de chacun des taxons (**Figure 10**), autrement dit de connaître le nombre de mailles 500 occupées pour chaque espèce, à l'aide de l'outil *Summary Statistics*. Une table contenant les fréquences de chacun des taxons pour l'intervalle de temps [2006-2017] au maillage 500 est obtenue à la fin de ce processus. Une variante de cette étape a été créée par la suite en sélectionnant uniquement les données dont la précision était égale ou inférieure à 250 m ou comprises entre 250 m et 1000 m.

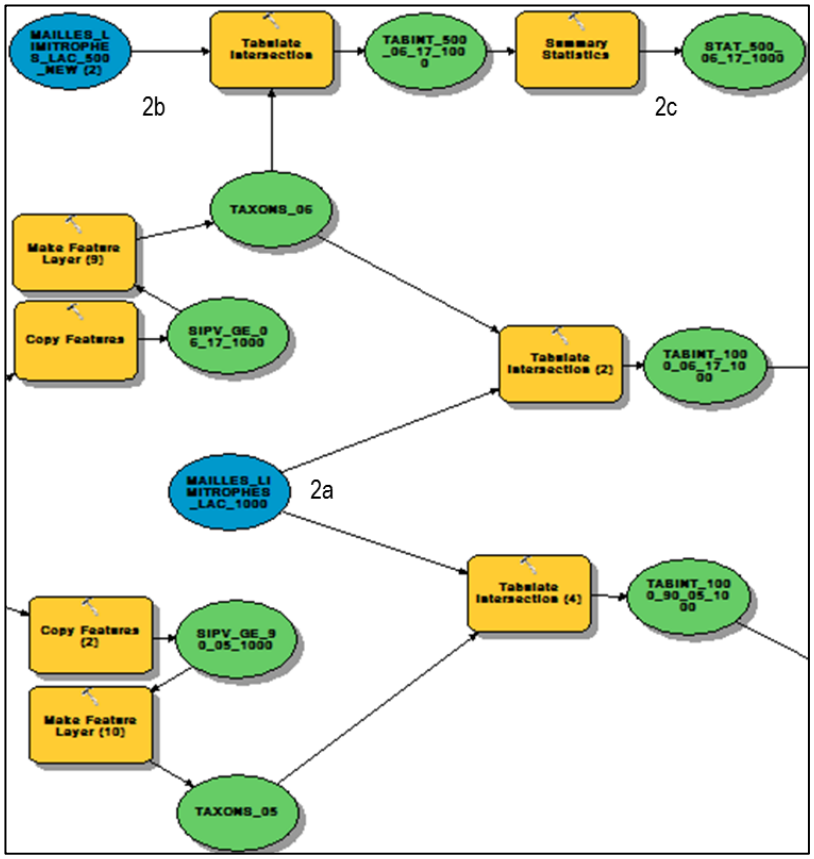


Figure 10 : Calcul des fréquences (nombre de mailles occupées) de chacun des taxons pour les deux pas de temps et les deux types de mailles.

La troisième étape du *ModelBuilder* est très semblable à ce qui vient être décrit ci-dessus et le but final est identique, à savoir la génération des fréquences pour chacun des taxons pour les deux intervalles de temps. Cependant, de nouveaux outils apparaissent. Dès lors, une brève description de leur utilisation à travers cette nouvelle étape me semble nécessaire. Par souci de simplification, seuls les processus pour l'intervalle de temps [2006-2017] sont décrits mais l'ensemble des opérations est effectué en parallèle pour le deuxième pas de temps.

A nouveau, l'outil *Summary Statistics* est utilisé afin d'établir la fréquence de chacun des taxons pour l'intervalle de temps considéré. L'outil *Alter Field* permet de renommer le champ COUNT_COORGE_ID issu du *Summary Statistics* par AOO1000_ ACTU. L'outil *Make Table View* autorise par la suite l'usage du *Table to Table* qui fonctionne de manière similaire à un *Copy Feature* mais cette fois-ci pour les tables. L'outil *Add Join* est ensuite utilisé afin de joindre à nouveau l'ensemble des identifiants genevois (NO_NOM_GE) présents dans la table **SIPV_1000_90_GE** (988 taxons) à la table **TAB_1000_06_17_1000** issue de la *Tabulate Intersection.* L'ajout de ce champ permettra aussi d'effectuer la jointure entre les deux tables des deux intervalles de temps. A l'issue de cette jointure, certains des taxons ont leur champ AOO1000_ACTU renseigné par un NULL car ces espèces ne sont pas présentes pour l'intervalle de temps. Les champs inutiles à la suite des traitements géomatiques sont ensuite éliminés avec l'outil *Delete Field*. Finalement, les NULL sont remplacés par des 0 à l'aide du *Calculate Field* (**Figure 11**) dans l'idée de pouvoir calculer l'évolution de la fréquence pour chacun des taxons entre les deux intervalles de temps.

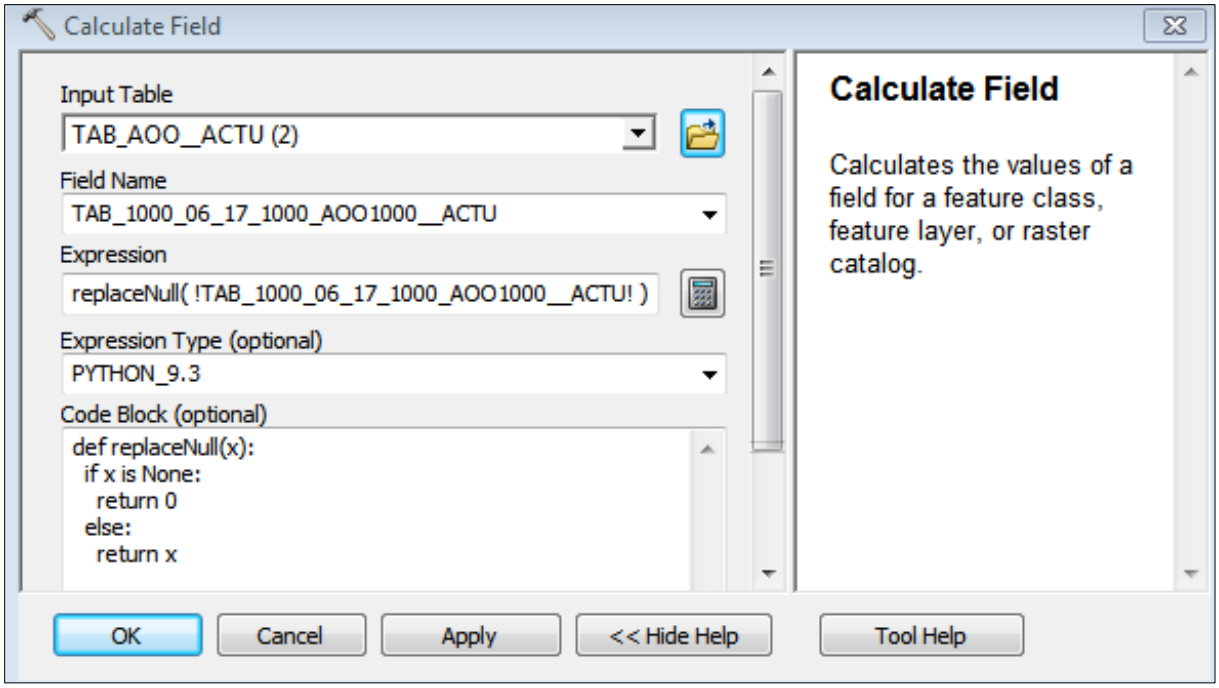


Figure 11 : Utilisation de l'outil *Calculate Field* pour remplacer les champs NULL par des zéros.

Les tables des deux intervalles de temps sont ensuite jointes via un *Join Field* sur l'attribut commun NO_NOM_GE.

La quatrième et dernière partie du *ModelBuilder* N°2 a pour but la génération et le calcul des différents champs qui permettront de calculer l'évolution de la fréquence pour chacun des taxons entre les deux intervalles de temps. Un premier champ, DELTA_NBRE_MAILLES, est créé et renseigné avec la

différence entre les fréquences de chaque taxon entre les deux l'intervalle de temps [1990-2005] et [2006-2017]. Un deuxième champ, EVOL_NB_MAILLES_1000, est créé afin de permettre le calcul de l'évolution, en pourcent, de la fréquence entre les deux pas de temps en remplissant les champs comme indiqué dans la **Figure 12** ci-dessous.

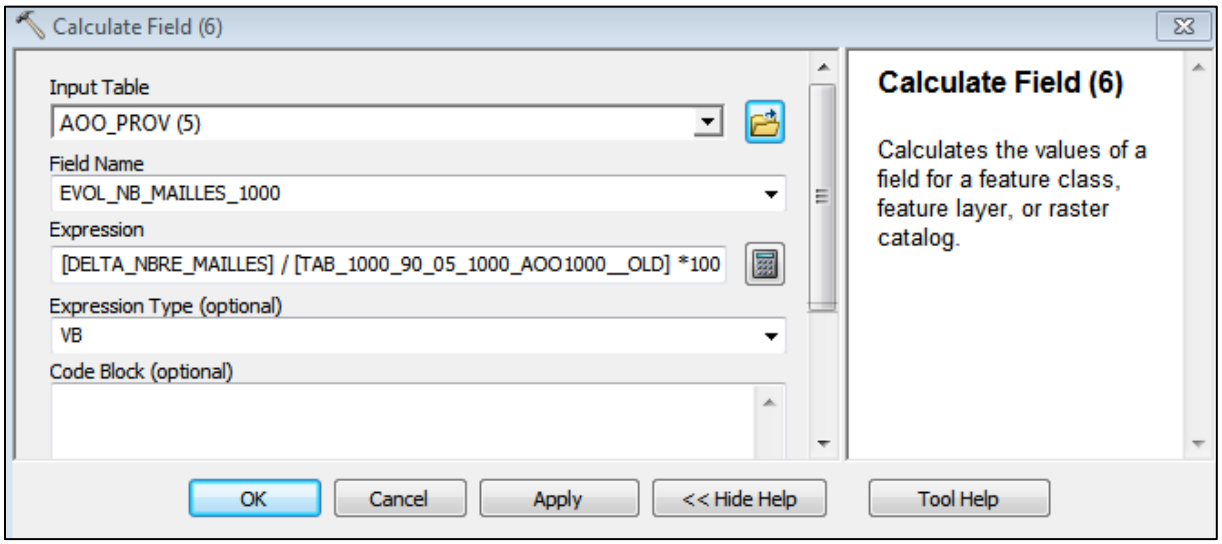


Figure 12 : Calcul de l'évolution du nombre de mailles, autrement-dit d'occurrences, entre les deux intervalles de temps.

Certaines espèces n'étaient pas prises en compte lors de la première liste rouge à l'instar d'*Himantoglossum robertianum* (Loisel.) P.Delfroge apparue récemment sur le territoire genevois. Ainsi, le calcul de l'évolution du nombre de mailles retourne un NULL, le dénominateur étant égal à zéro. Le problème est contourné en remplaçant les NULL par 1000 en remplissant les champs de l'outil *Calculate Field* comme dans la **Figure 13** ci-dessous.

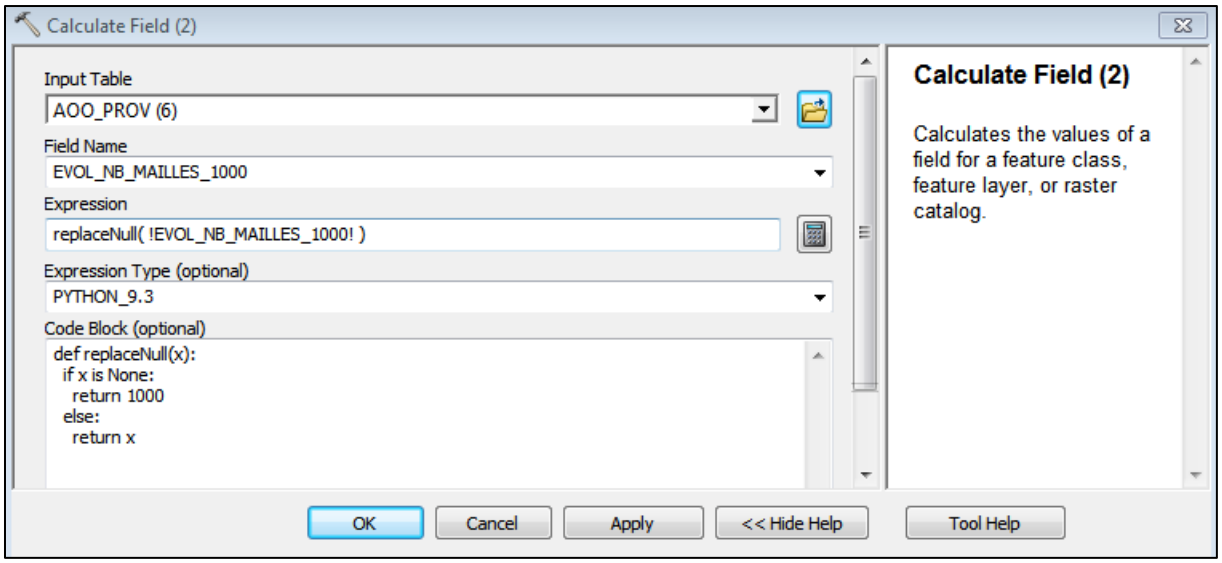


Figure 13 : Affranchissement des divisions par 0 où les NULL résultant sont remplacés par 1000.

La table finale, **AOO_PROV_FIN**, est ensuite exportée dans la gdb **PROD_OUTPUT.gdb** ou sous la forme d'un fichier Excel.

3.7 Attribution des statuts de menaces provisoires selon le critère A2c

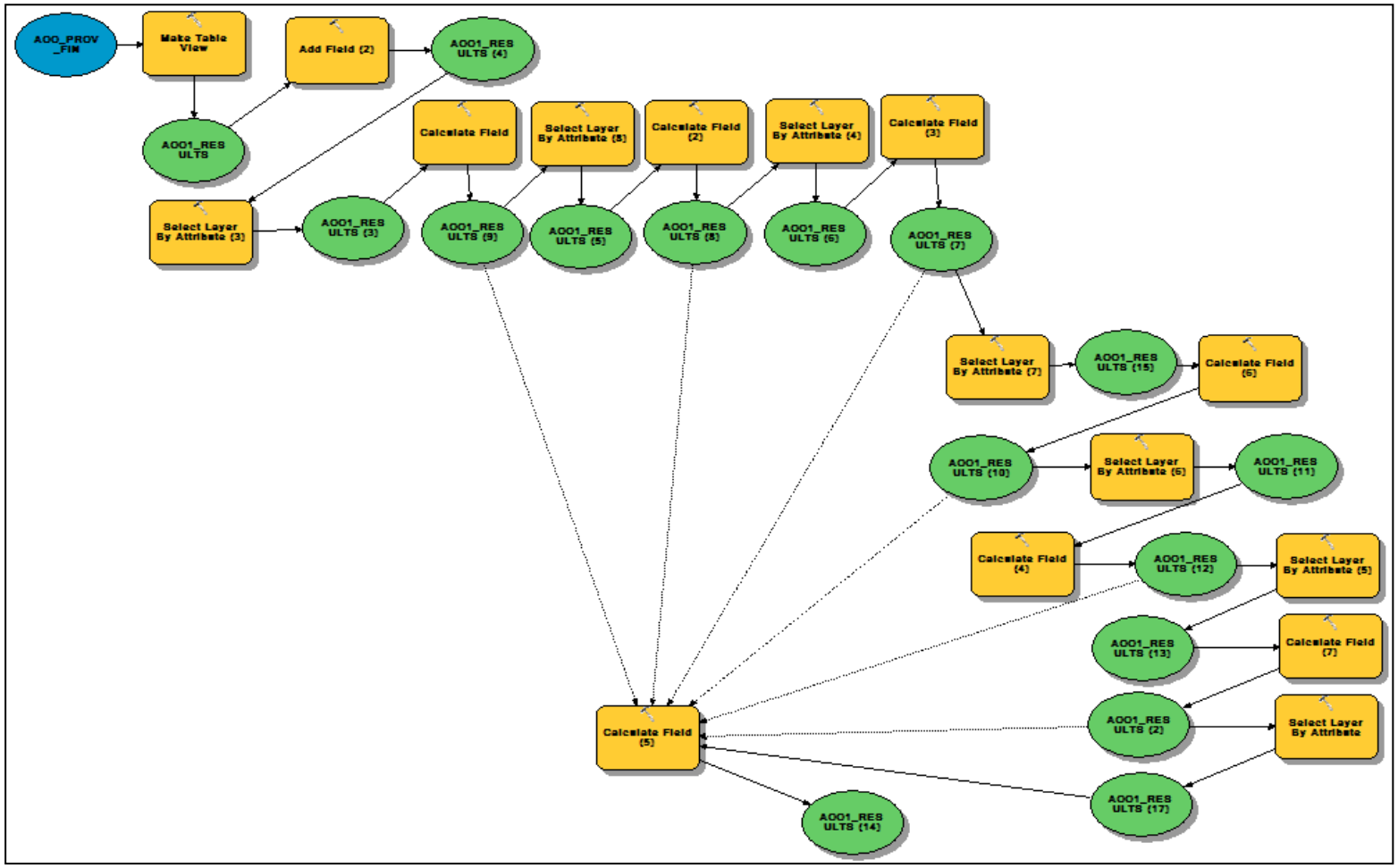


Figure 13 : Attribution des statuts de menaces provisoires selon le critère A2c régionalisés. Les préconditions sont représentées par les traits tillés.

Le but de ce troisième *ModelBuilder* est d'attribuer un statut de menace basé sur le pourcentage de régression (ou d'augmentation) des fréquences entre les deux pas de temps calculé dans le *ModelBuilder* N°2 et disponible dans la table **AOO_PROV_FIN**. Le modèle est basé sur une suite de sélections par attribut suivies de calculs de champs permettant l'attribution des statuts de menace. Toute l'astuce de ce modèle repose sur l'utilisation de préconditions qui permettent de renseigner le champ STATUT_PROVISOIRE au fur et à mesure des sélections et des calculs de champs (**Figure 13**).

Prenons l'exemple d'*Aceras anthropophorum* (L.)W.T.Aiton afin de mieux cerner le fonctionnement du modèle. Par le *ModelBuilder* N°2, nous savons que l'espèce était présente dans 46 mailles kilométriques dans le premier intervalle de temps mais plus que dans 43 pour le second. Le modèle nous indique ainsi que la régression du nombre d'occurrences est de 6.52 % entre les deux pas de temps. A partir de ce résultat, le *Model Builder* N°3 assigne un 5 comme statut de menace provisoire à l'espèce. Selon le critère A, l'espèce serait ainsi quasi menacée. La **Figure 14** tente une synthèse du fonctionnement du *Model Builder N°3* pour l'exemple *d'Aceras anthropophorum* (L.)W.T.Aiton. Les résultats provisoires sont donnés dans l'**Annexe 2**.

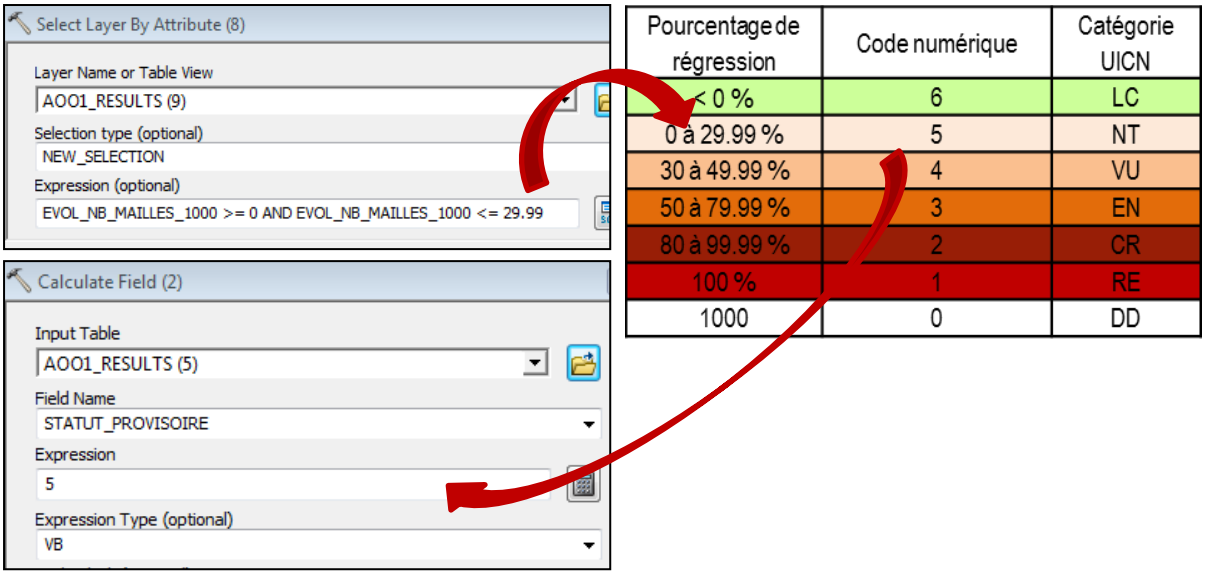


Figure 14 : Synthèse du fonctionnement de l'attribution des statuts de menace provisoire selon le critère A2c.

3.8 Critère B2. Répartition géographique et créations des localités

3.8.1 Génération des localités

Plusieurs modèles ont été testés pour la génération des localités. Quatre d'entre eux sont présentés dans le présent document. Aucun ne donne de solutions définitives à la génération de localités mais seulement un faisceau de pistes plus au moins abouties.

3.8.2 Définition d'une localité

Concrètement, une localité au sens de la démarche d'une liste rouge, *définit une zone particulière du point de vue écologique et géographique dans laquelle un seul phénomène menaçant peut affecter rapidement tous les individus du taxon présent. L'étendue de la localité dépend de la superficie couverte par le phénomène menaçant et peut inclure une partie d'une sous-population au moins. Lorsqu'un taxon est affecté par un phénomène menaçant au moins, la localité doit être définie en tenant compte de la menace plausible la plus grave.* **(UICN, 2011)**. Il est possible d'illustrer cette définition par l'exemple fictif suivant. Imaginons-nous sur la rive Sud du lac de Neuchâtel en plein cœur des réserves de la Grande Cariçaie. Dans chacune des 8 réserves croissent des *Viola persicifolia* Schreb.. Il existe donc 8 stations de *V. persicifolia* mais seulement une localité de cette espèce qui serait l'ensemble de la rive Sud du lac de Neuchâtel. En effet, toutes les stations se trouvent dans les réserves et sont soumises au même mode de gestion, soit une fauche quadriennale pour éviter l'embuissonnement des

prairies à choin (*Orchio-Schoenetum nigricantis*) où croît l'espèce. Le phénomène menaçant commun à toutes ces stations pourrait être une crue du lac de Neuchâtel qui noierait l'ensemble des stations de *V. persicifolia*. Ce phénomène menaçant toucherait l'ensemble des stations de façon rapide et simultanée. C'est d'ailleurs ce qui était arrivé lors de la crue de mai 2016 où le lac était monté de 80 cm noyant l'ensemble des populations de violettes à feuilles de pêcher. Cela a permis de pêcher des brochets dans les aulnaies. Fin de la parenthèse.

En fait, l'automatisation de la création des cartes de localités par espèce est relativement complexe, voire quasi impossible car la notion même de menace peut être subjective. Ainsi, regrouper l'ensemble des stations de *Scrophularia canina* L. du vallon de l'Allondon sous une même localité est intuitif et facile. L'idée d'une crue dévastatrice qui emporterait l'ensemble des stations bordant la rivière est facilement imaginable et le vallon est une entité facilement délimitable géographiquement. L'exercice devient plus difficile et nettement moins intuitif avec quelques pieds d'*Asplenium trichomanes* L. croissant sur un muret long d'une centaine de mètres et qui appartiendrait à plusieurs propriétaires. Répondre qu'il n'y aurait qu'une localité pourrait s'avérer tout simplement faux dès lors qu'il existerait des propriétaires plus enclins que d'autres à passer au Karcher annuellement leur portion de mur afin de respecter le propre en ordre helvétique. La menace n'est donc pas homogène sur l'ensemble du muret d'où la création de plusieurs localités sur une centaine de mètres. Par ces quelques exemples triviaux, il devient aisé de saisir que la création de localités automatisée est une gageure. Les différents modèles présentés ci-dessous se rapprochent cependant d'une semi-automatisation de la génération des localités par espèces.

3.8.3 *ModelBuilder* pour la création des prélocalités

Les localités ne sont calculées que pour les espèces dont les occurrences sont inférieures ou égales à 60 mailles 500 et qui présentes une diminution du nombre de mailles kilométriques entre les deux intervalles de temps.

Modèle N°1

Les premières étapes sont effectuées en dehors du *ModelBuilder* car il arrive fréquemment que le modèle « tourne dans le vide » lors de la première étape de constitution des *Buffers*. Ce problème est d'autant plus fréquent que le modèle est *runné* plusieurs fois de suite. Pour contourner ce problème, une table contenant les taxons présents dans moins de 61 mailles de 500 m et dont la précision est égale ou inférieure à 250 m est créée. Cette table découle directement de celle obtenue précédemment dans l'étape de création des AOO avec le maillage de 500 m du *ModelBuilder* N°2. Les données sont ainsi celles du deuxième intervalle de temps. Un premier *Buffer* est effectué sur la précision des notes floristiques (XY_PRECISION). La table **T8_SIPV_250_06_17_1000_GE_NET_AOO_61_BUFF** est ainsi créée. Parallèlement, une table vide mais contenant les mêmes champs que la table **T8_SIPV_250_06_17_1000_GE_NET_AOO_61_BUFF** est créée. La table peut ensuite être introduite dans le modèle qui peut être grossièrement séparé en deux sous-modèles. Il est souhaitable de faire tourner ces deux parties l'une après l'autre.

La première partie du modèle (**Figure 16**) se base sur la précision des notes floristiques, représentées par les *Buffers* établis précédemment, pour la création des localités. La première partie du modèle sépare les données en deux groupes à l'aide de sélections par attribut [1] et [2]. Le premier groupe contient les données dont la précision est inférieure ou égale à 25 mètres (XY_PRECISION <= 25) et le

second regroupe les données dont la précision est supérieure à 25 mètres (XY_PRECISION > 25). Une sélection spatiale opérée à l'aide de l'outil *Select Layer By Locations* [3] est ensuite effectuée sur les deux sélections précédentes (**Figure 15**).

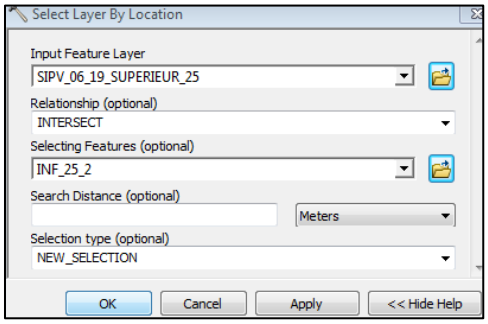


Figure 15 : Sélection par emplacement de notes dont la précision est supérieure à 25 et qui intersectent des notes plus précises de la même espèce.

Les notes dont la précision est supérieure à 25 mètres et qui *intersectent* des notes plus précises sont inscrites dans la table **Aceras_poubelle** à l'aide de l'outil *Append* [4] et seront par la suite éliminées lors des traitements géomatiques futurs. En revanche, si une note a une précision supérieure à 25 mètres mais qu'elle n'en *intersecte* aucune autre, alors elle sera conservée. Une copie des notes éliminées et inscrites dans la table **Aceras_poubelle** est finalement effectuée (**POUBELLES_RESULTATS**). L'itération effectuée à l'aide de l'*Iterate Row Selection* [5] permet de répéter l'ensemble de ce processus sur l'ensemble des taxons. Ainsi, sur les 25 123 notes (issues du deuxième intervalle de temps et dont la précision est égale ou inférieure à 250 m) introduites en entrée du modèle, 911 sont inscrites dans les notes à éliminer (**Figure 16**).

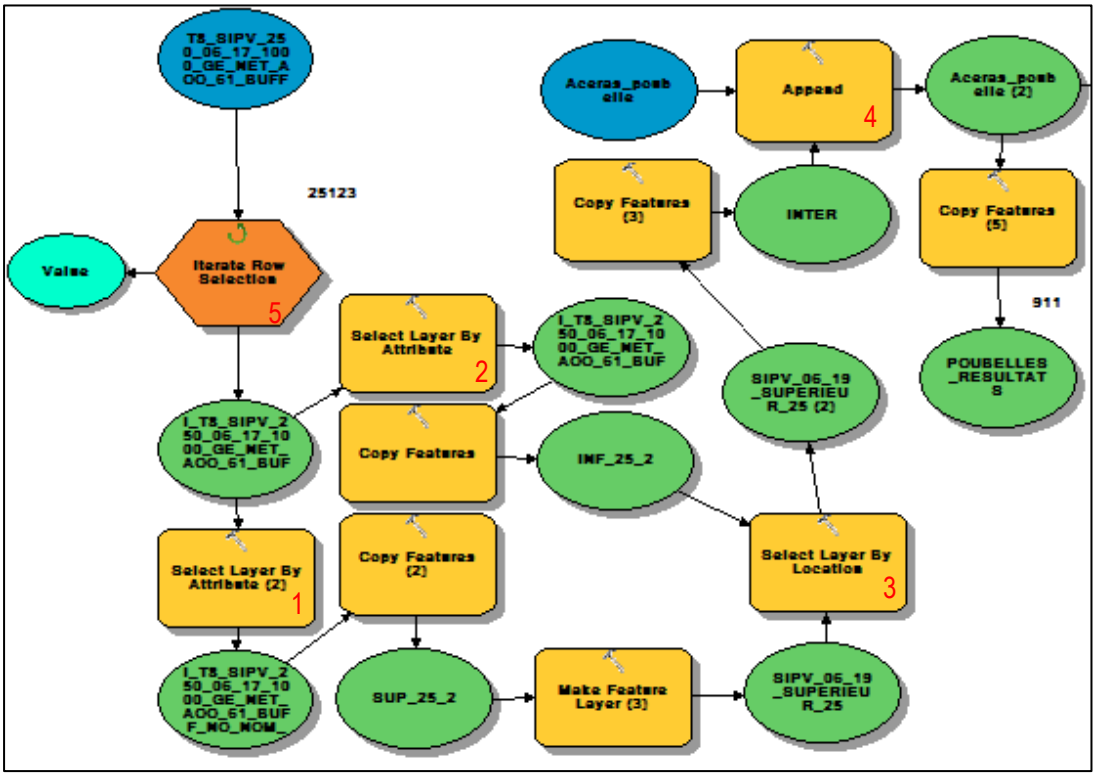


Figure 16 : Première partie du modèle de la méthode N°1 permettant de séparer les notes précises de celles qui le sont moins.

La deuxième partie du modèle consiste à éliminer les données peu précises et à fusionner les notes restantes sur la base de leur *Buffer* afin de créer les prélocalités pour chacune des espèces. Un nouveau champ, PRECISION, est ajouté à la table **Aceras_poubelle**. A l'aide de l'outil *Calculate Field*, le champ est renseigné avec la valeur 2. Un *Add Join* basé sur le champ ORIG_FID permet ensuite de joindre la table **T8_SIPV_250_06_17_1000_GE_NET_AOO_61_BUFF** à **Aceras_poubelle**. Un contrôle du nombre de données nous indique que nous avons à nouveau 25 123 notes. Une sélection par attribut (Aceras_poubelle_PRECISION IS NULL) permet ensuite de sélectionner les notes pour lequel le champ PRECISION n'a pas encore été renseigné, autrement dit l'ensemble des données que nous cherchons à garder. Le champ est alors renseigné par un 1. Une nouvelle sélection par attribut sur le champ PRECISION (Aceras_poubelle_PRECISION <> 2) suivi d'une copie (**SIPV_250_06_17_BUFF_MAX_PRECISION**) de la sélection permet de ne garder que les notes les plus précises. Un rapide contrôle du nombre de notes nous indique que la copie ne contient plus que 24 212 données. Les 911 données peu précises ont donc bien été éliminées. L'ensemble des champs provenant de la table **Aceras_poubelle** sont ensuite supprimés. Finalement, l'outil *Dissolve* avec l'option *Create multipart features* décochée permet de regrouper via leur identifiant genevois les *Buffers* d'une même espèce et d'esquisser ainsi leur localité. Un *Summary Statistics* permet ensuite de connaître le nombre des prélocalités ainsi constituées par taxon.

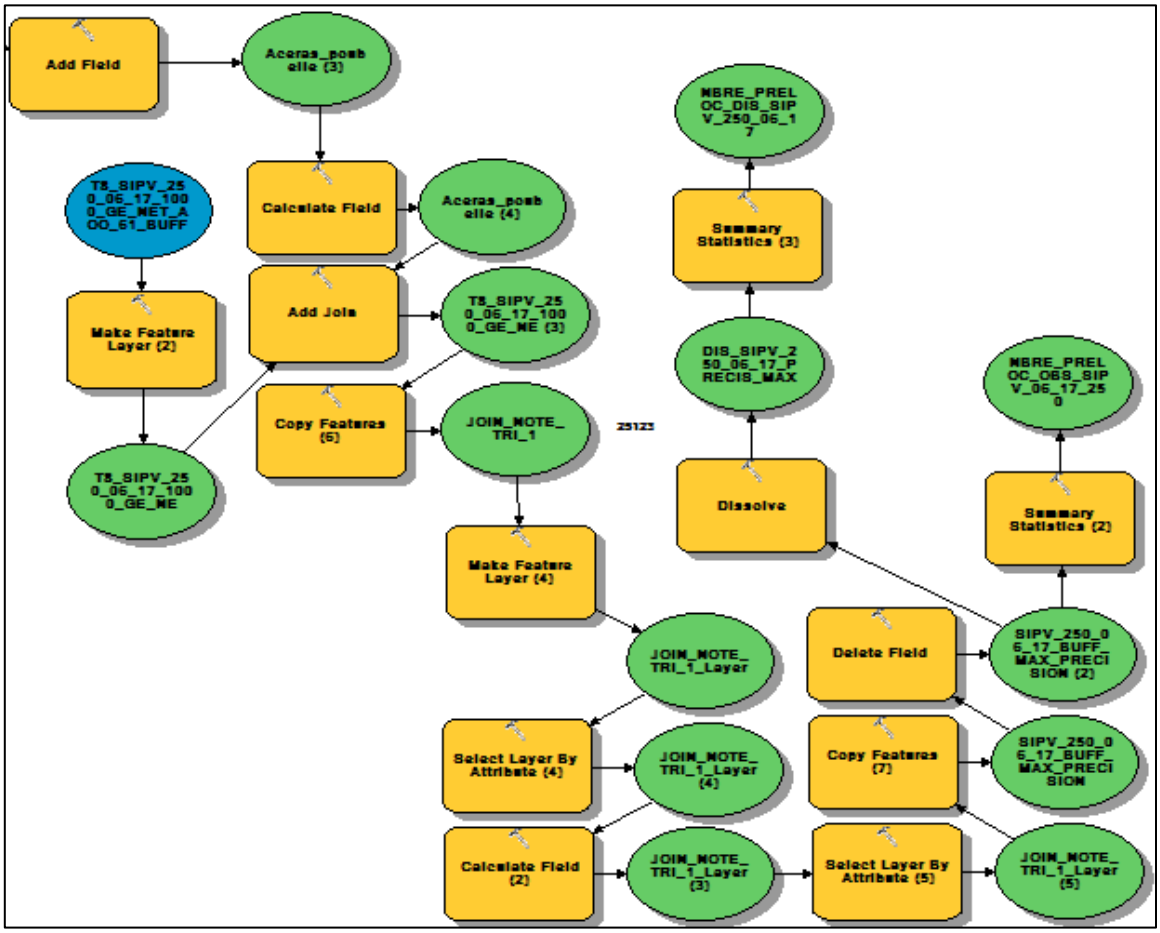
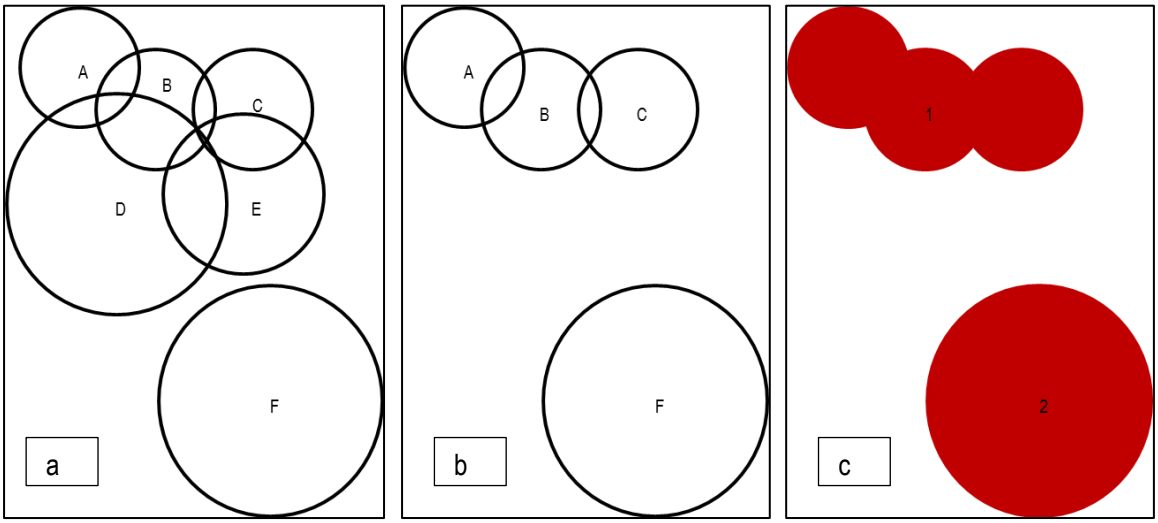


Figure 17 : Deuxième partie du modèle de la méthode N°1 permettant le calcul l'élimination des notes peu précises et le calcul des prélocalités

Les illustrations (**Figures 18a**, **18b**, **18c**) ci-dessous résument l'ensemble des opérations effectuées au sein du *ModelBuilder*.



Figures 18a, 18b, 18c : **18a** représente l'état initiale. **18b** les notes restantes après élimination des notes dont la précision est supérieure à 25 m et qui *intersectent* des notes plus précises. **18c** indique les deux prélocalités obtenues après la fusion des *Buffers* des notes sauvegardées.

A, B, C, D, E, F représentent les *Buffers* d'une espèce lambda obtenus à partir du champ XY_PRECISION. Le premier sous-modèle élimine les notes D et E à la fois parce que le rayon de leur *Buffer* respectif est plus grand que 25 m mais aussi parce qu'elles *intersectent* des notes plus précises. Ainsi, la note F, bien qu'étant la plus imprécise des six, n'est pas éliminée car elle n'*intersecte* aucune autre note dont la précision serait égale ou inférieure à 25 mètres. Les *Buffers* qui n'ont pas été éliminés sont ensuite fusionnés et de cette fusion découle la création de deux prélocalités.

Cette méthode présente un inconvénient majeur car elle a tendance à multiplier le nombre de prélocalités de façon artificielle et d'éliminer des notes très précises. Le schéma ci-dessous (**Figure 19**) devrait aider le lecteur à prendre conscience de cet écueil.

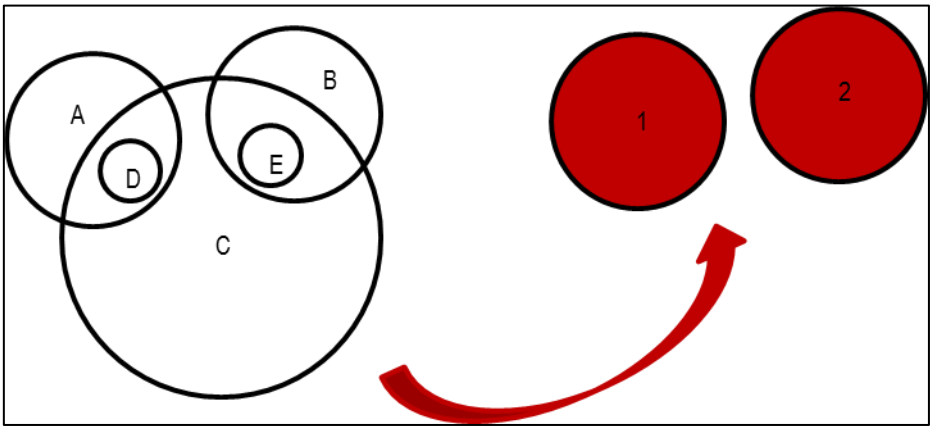


Figure 19 : Les notes D et E sont fusionnées aux notes A et B lors de la constitution des prélocalités alors qu'elles sont pourtant plus précises. Les deux prélocalités sont en outre spatialement très proches l'une de l'autre.

Les notes A et B sont spatialement très proches mais forment deux localités distinctes (1 et 2) après élimination de la note C dont la précision et moindre. De plus, les notes D et E disparaissent lors de l'étape de fusion des *Buffers* et cela malgré une précision plus fine que les données A et B.

MéthodeN°2

La méthode N°2 standardise la précision des notes à 25 mètres. Autrement dit, un *Buffer* de 25 m de rayon est créé pour l'ensemble des notes. Ainsi, certaines notes perdront de leur précision, alors que d'autres deviendront artificiellement plus précises (**Figure 20**).

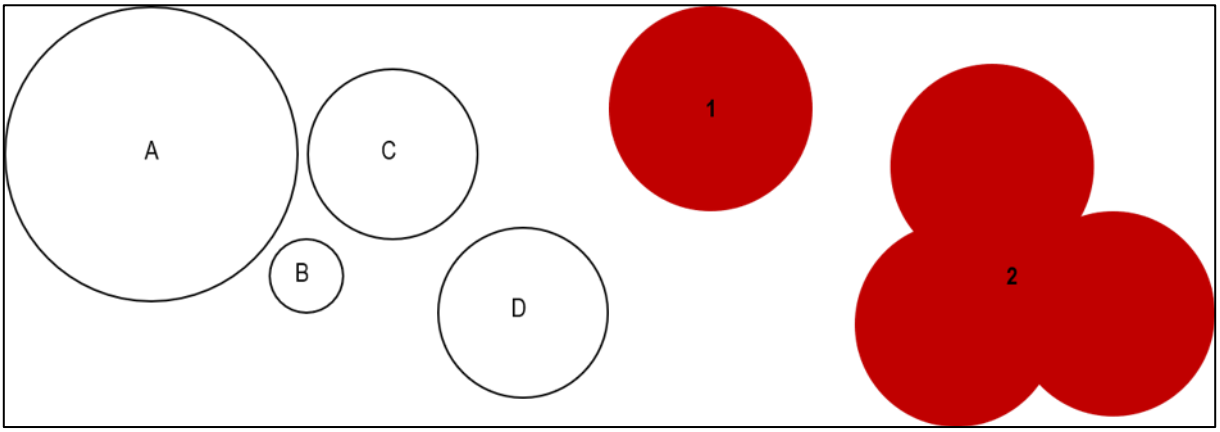


Figure 20 : Standardisation (25 m de rayon) et dissolution des zones tampons pour la construction de deux prélocalités.

Une fois la standardisation effectuée, les *Buffers* sont dissouts afin de créer les localités. Comme pour la méthode N°1, une multitude de prélocalités sont créées mais celles-ci ne correspondent pas forcément à définition même de la localité.

Le *ModelBuilder* (**Figure 21**) utilisé pour la création des prélocalités est simplissime et reprend bon nombre d'outils utilisés dans la méthode N°1. Dès lors, je ne le décrirai que de manière sommaire.

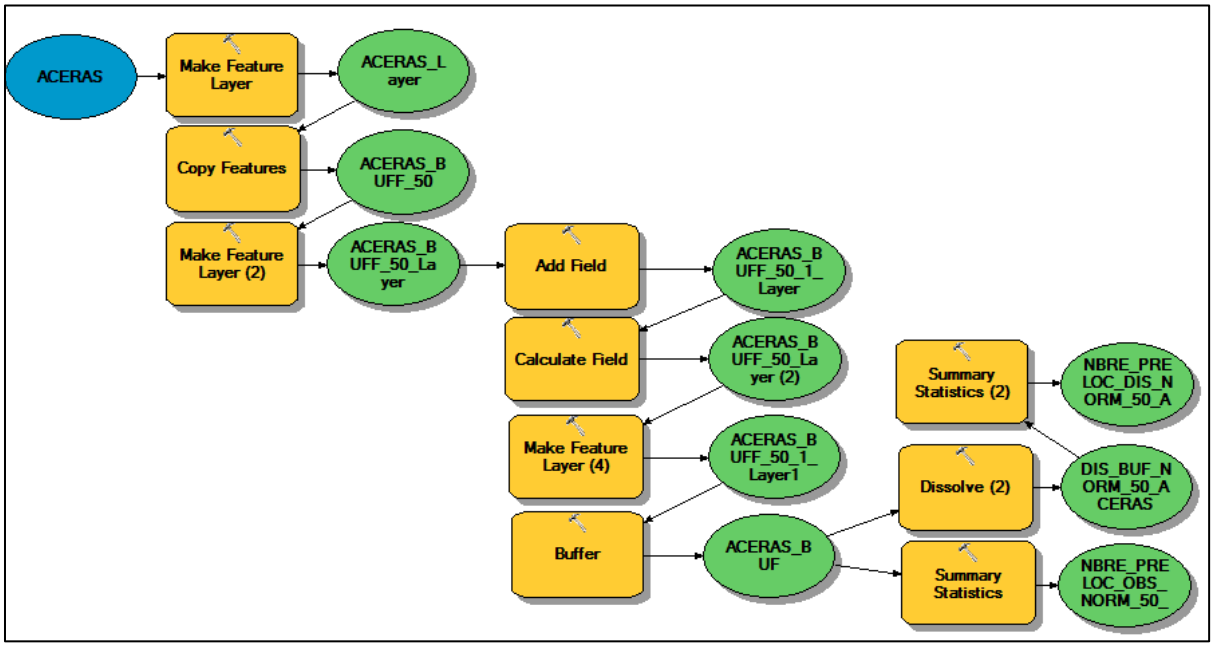


Figure 21 : *ModelBuilder* de la méthode N°2.

La table **ACERAS** avait été utilisée pour tester le modèle sur un seul taxon, mais le modèle a néanmoins été testé sur l'ensemble des données de l'intervalle de temps [2006-2017], dont la précision est égale ou inférieure à 250 m et le nombre d'occurrences basé sur les mailles de 500 m inférieur à 61. Comme pour les autres modèles, une copie de la table est effectuée en début du processus afin de

récupérer facilement d'éventuelles erreurs, mais aussi de stabiliser le modèle. Un champ, BUF_50, est ajouté à la copie **ACERAS_BUFF_50** via l'outil *Add Field* puis renseigné avec la valeur 25 à l'aide du calculateur de champ. Les *Buffers* sont ensuite créés sur la base de ce nouveau champ. Finalement, de façon identique au *Model Builder* de la méthode N°1, un *Summary Statistics* est effectué afin de connaître le nombre de prélocalités avant et après la dissolution sur le champ NO_NOM_GE des *Buffers* par l'outil *Dissolve* dont l'option *Create multipart features* a été préalablement décochée.

Il arrive fréquemment que les *Buffers* soient créés selon la précision originale de la note et non sur la standardisation à 25 m lors de l'utilisation du *Model Builder*. Ce problème peut être contourné en supprimant manuellement le champ XY_PRECISION de la table introduite en entrée du modèle.

Méthode N°3

La méthode N°3 est complexe car elle fait intervenir plusieurs étapes de préparation des données plus en amont et un modèle est créé pour chacun des onze groupements de milieux naturels.

Un ou plusieurs groupements de milieux naturels sont attribués à chacune des espèces considérées au sein de la nouvelle liste rouge. Ainsi, onze tables, **Tab_GRPT_XX**, sont créées pour chacun des groupements de milieux naturels. A chaque fois, table contient les espèces susceptibles d'être rencontrées au sein du groupement de milieux naturels. Les taxons ont été attribués de la façon suivante aux différentes tables. Chaque espèce possède un code Delarze indiquant le milieu naturel auquel elle est rattachée, ou plutôt pour lequel elle est caractéristique. Une correspondance est par la suite effectuée entre les milieux naturels établis par Delarze et les milieux naturels genevois. Les espèces possèdent un code milieu, CO_MILIEU, une fois la correspondance effectuée ce qui permet par la suite de les relier à un milieu naturel genevois et de là à un des onze groupements. Une table est alors créée par groupement. Certaines espèces sont caractéristiques de milieux naturels qui n'existent pas à Genève. Il faut alors associer manuellement les taxons concernés à un milieu naturel genevois permettant ainsi leur assignation au sein d'un groupement. Au final, les tables de groupements devraient posséder les champs suivants ; OBJECTID, NO_NOM_GE, CO_MILIEU et NO_GRPT.

Les figures suivantes détaillent le *Model Builder* de la méthode N°3 et se focalisent sur le groupement N° 9 (Forêts et ourlets mésophiles). Comme pour les deux modèles précédents, seules les données dont la précision était inférieure ou égale à 250 ont été prises en compte et le nombre d'occurrences par taxon s'est basé sur le maillage de 500 m. La première étape consiste à joindre [1] à la table **Tab_GRPT_09** la table **AOO_61** sur la base du champ NO_NOM_GE (**Figure 22**). La table **AOO_61** contient 674 taxons dont le nombre d'occurrence est inférieure à 61 pour l'intervalle de temps [2006-2017]. Lors de la jointure, l'ensemble des taxons qui n'appartiennent pas au groupement N°9 auront un champ NO_NOM_GE renseigné par un NULL. Ces taxons sont alors éliminés à l'aide d'une sélection par attribut (AOO_61.NO_NOM_GE IS NOT NULL) [2] permettant de ne garder que les espèces liées au groupement N°9. La table ne contient cependant aucune information quant à la précision de la note, ses coordonnées géographiques ou bien encore son nom latin. La jointure de champ (*Join Field*) [3] entre la table **Tab_GRPT_09** et la table **T8_SIPV_250_06_17_1000_GE_NET_AOO_61** basé sur le champ NO_NOM_GE permet de les récupérer. Une nouvelle sélection par attribut [4] permet d'éliminer à nouveau les espèces qui n'ont pas été attribuées au groupement (AOO_61_NO_NOM_GE IS NOT NULL).

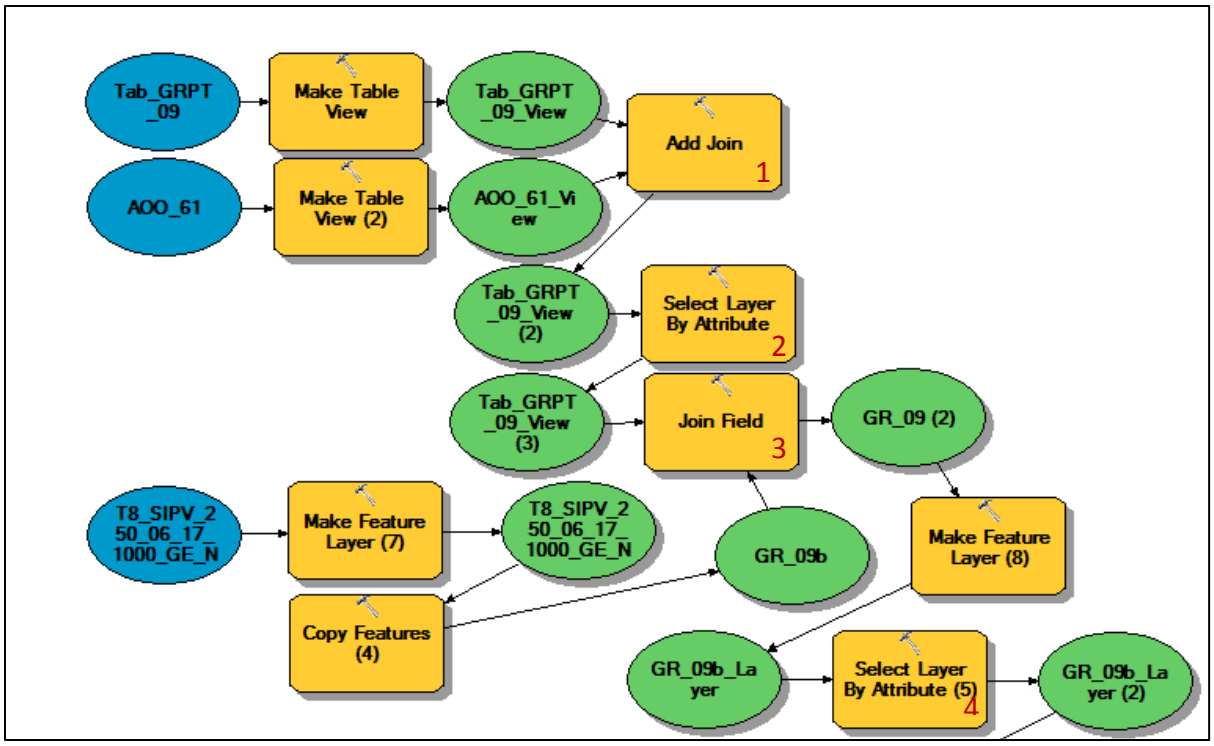


Figure 22 : Jointure de la table du groupement de milieux naturel 09 à la table des taxons retenus.

Les champs inutiles sont ensuite éliminés via l'outil *Delete Field* [5] (**Figure 23**). A la suite de cette opération, il ne devrait rester que les champs suivants : NO_NOM_IF, NOM_COMPLET, A_NOTE, X, Y, XY_PRESICION, NO_NOM_GE.

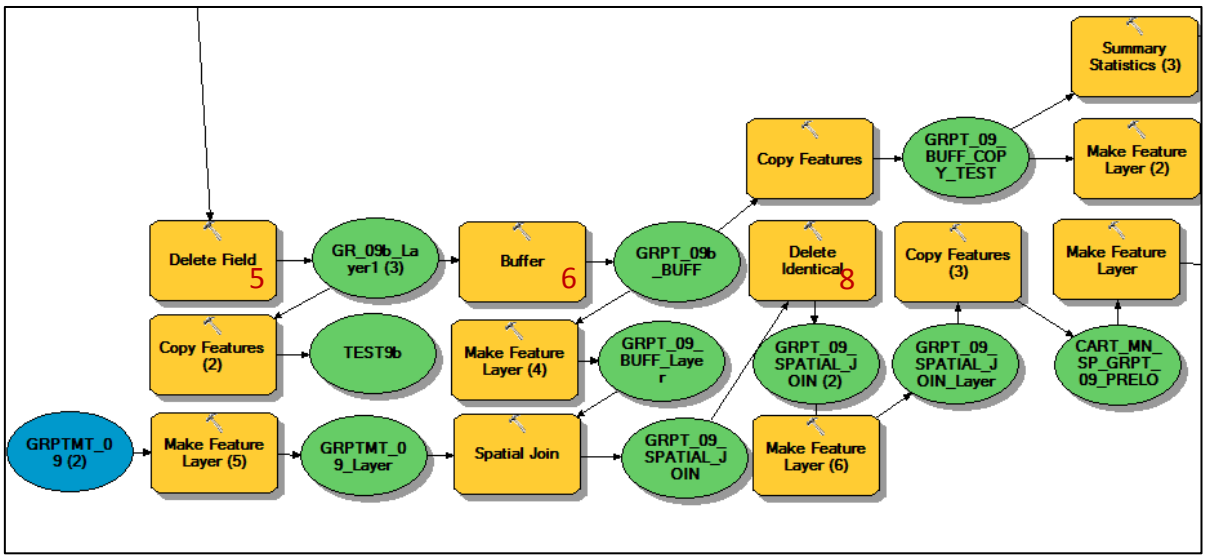


Figure 23 : Elimination des champs inutiles et création des *Buffers* basés sur la précision des notes floristiques.

A nouveau, l'outil Buffer [6] est appliqué sur le champ PRECISION_XY afin de tracer l'imprécision de chacune des notes floristiques. Une jointure spatiale est ensuite effectuée à l'aide de l'outil Spatial Join *[7]* entre la couche **GRPTMT_09** et les Buffers des différentes espèces obtenus à l'étape précédente. Les polygones du groupement N°9 qui touchent un *Buffer* d'une espèce s'allument (**Figure 24**) et délimitant ainsi une prélocalité.

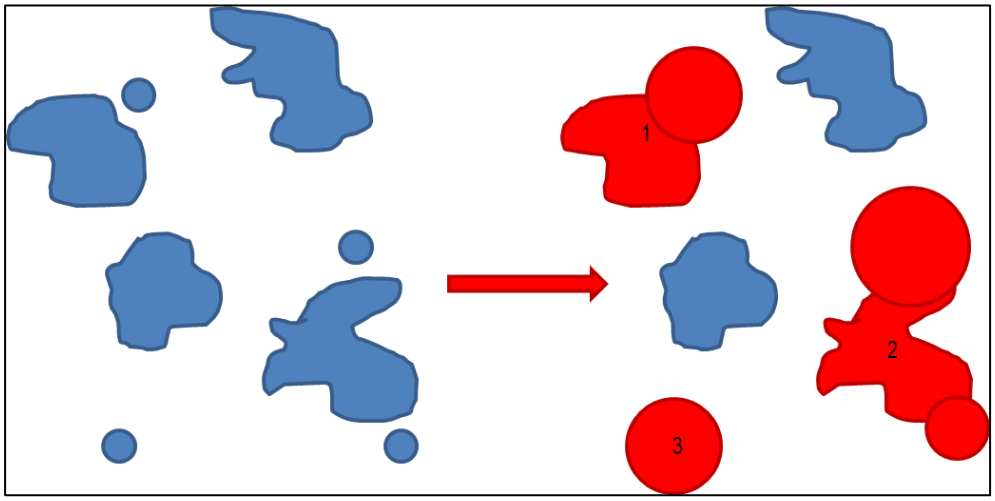


Figure 24 : Les *Buffers* qui touchent un groupement de milieux naturels favorables les allument délimitant ainsi des prélocalités (en rouge).

Il arrive fréquemment que plusieurs *Buffers* d'une même espèce touchent un polygone du groupement de milieux naturels ce qui a pour effet de créer des doublons au sein de la table **GRPT_09_SPATIAL_JOIN** obtenue lors de la jointure spatiale. L'outil *Delete Identical* [8] appliqué conjointement aux champs TARGET_FID et NO _NOM_IF permet d'y remédier. Une copie de la couche obtenue (**CART_MN_SP_GRPT_09_PRELOC_TEST**) est ensuite effectuée [9]. Celle-ci permettra par la suite de générer des cartes de localités basées sur les groupements de milieux naturels pour chacune des espèces du groupement N°9 à l'aide du Model Builder 9_MB_SP_MN_PRELOC. Puis un *Select Layer By Location* [11] combiné à un *Summary Statistics* [12] (**Figure 25**) nous permettent de connaître le nombre de notes qui ont été attribuées au groupement N°9.

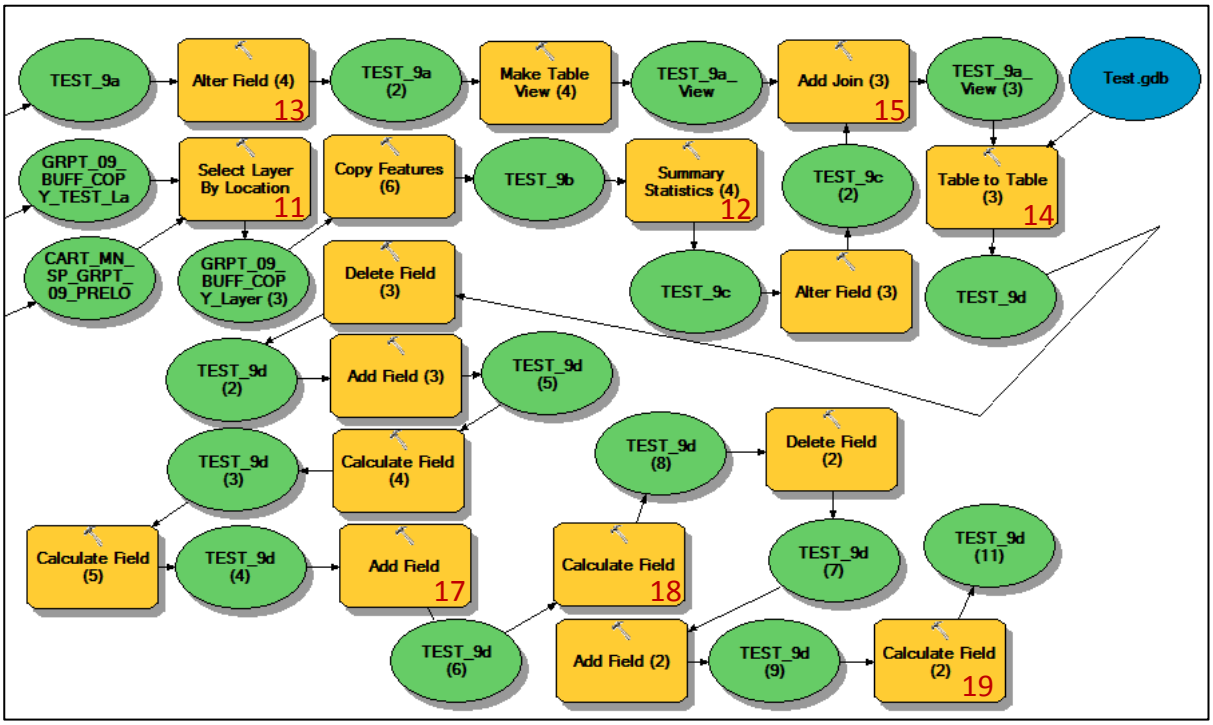


Figure 25 : Décompte du nombre de prélocalités obtenues à l'aide du modèle de la méthode N°3.

Parallèlement, le nombre de *Buffers*, autrement dit le nombre de notes par espèce, est décompté via un *Summary Statistics* [10]. Le champ FREQUENCY résultant de l'opération est ensuite renommé

NBRE_BUF_TOT [13]. Les deux tables sont ensuite jointes et un *Table to Table* [14] permet une vérification des différents processus tel que le renommage des champs, mais aussi de copier au sein d'une gdb le résultat de la jointure [15]. Un nouveau champ, DIFF_BUFF_GRPMT [16], est ensuite ajouté à la table et renseigné avec la différence entre le nombre de *Buffers* total, NBRE_BUF_TOT, et ceux qui *intersectent* le groupement, BUF_INCLUS_GRPMT. Un troisième champ, CO_GRPMT, est ensuite ajouté [17] et renseigné avec le numéro du groupement à l'aide du calculateur de champ [18]. Ce nouveau champ permettra de ne pas perdre les informations contenues dans chacun des tableaux des modèles de la méthode N°3 une fois lorsqu'ils seront réunis au sein de la table de synthèse finale. La dernière partie du modèle [19] permet de calculer la fidélité de l'espèce à son groupement de milieux naturels par la division suivante : [BUF_INCLUS_GRPMT] / [NBRE_BUF_TOT] *100. Le résultat final de l'ensemble de ces processus ne suffit cependant toujours pas à déterminer le nombre de localités par espèces. A ce stade, nous connaissons uniquement le nombre de notes qui sont incluses dans le bon groupement de milieux naturels et celles qui se trouvent en dehors. Nous pouvons aussi connaître le nombre de polygone de groupement de milieux naturels par espèce à l'aide d'un *Statistics Summary* effectué sur les couches **CART_MN_SP_GRPT_09_PRELOC_XX**.

Méthode N°4 (modèle 7_MB_GEN_LOC_09c)

La conception de ce modèle est très similaire à ce qui a déjà été fait dans les méthodes précédentes. La force de ce nouveau modèle réside dans le « repêchage » des notes qui tombent hors des groupements par la standardisation de leur précision à 50 m suivie d'une dissolution des *Buffers* nouvellement obtenus.

Les premières étapes étant pratiquement identiques à celles de la méthode N°3, il ne me semble pas nécessaire de les développer de façon détaillée. Le groupement N°10 est pris ici comme exemple, mais la conception de la table **Tab_GRPT_10** et de la couche **GRPMT_10** ne diffère aucunement de ce qui avait été déjà énoncé antérieurement pour la méthode N°3. A nouveau la table **AOO_61** est jointe à la table **Tab_GRPT_10** sur le champ NO_NOM_GE. Une fois la jointure effectuée, une partie du champ NO_NOM_GE est renseignée par des NULL car la table **Tab_GRPT_10** ne possède pas l'ensemble des 674 taxons contenus dans la table AOO_61. Une sélection par attribut permet d'éliminer les taxons dont le champ NO_NOM_GE est renseigné par un NULL. Une jointure de champs est effectuée entre les tables **Tab_GRPT_10** et **ACERAS** qui est une sélection des notes d'*Aceras anthropophorum* de la table **T8_SIPV_250_06_17_1000_GE_NET_AOO_61** sur la base du champ NO_NOM_GE. A nouveau, la correspondance n'est pas parfaite entre les deux tables. Une nouvelle sélection par attribut permet d'éliminer l'ensemble des notes dont le champ NO_NOM_GE est renseigné par un NULL. Dans le cas présent, cela revient à éliminer l'ensemble des taxons à l'exception de l'*Aceras anthropophorum*. L'outil *Delete Field* permet de supprimer l'ensemble des champs inutiles à la suite des traitements géomatiques, autrement dit les champs provenant des tables **Tab_GRPT_10** et **AOO_61** issus des jointures successives. Comme pour la méthode N°3, un *Buffer* est appliqué sur l'ensemble des notes sur la base du champ XY_PRECISION. Une jointure spatiale est ensuite effectuée entre les *Buffers* précédemment obtenus et la couche **GRPMT_10** sur laquelle une zone tampon de 10 mètres a été appliquée dans le but de récupérer un maximum de notes au sein du groupement. A nouveau, plusieurs *Buffers* d'une même espèce peuvent toucher un même polygone de groupement de milieux naturels ce qui a pour effet de créer des doublons au sein de la table obtenue suite à la jointure spatiale. L'outil *Delete Identical* appliqué conjointement sur les champs TARGET_FID et NO_NOM_IF permet de les

éliminer. Les notes qui ne sont pas prises en compte lors de la jointure spatiale sont alors récupérées par une première sélection par emplacement (*Select Layer by Location*) avec l'option *Switch Selection* sélectionnée, puis leur précision est standardisée à 50 m (25 m de rayon). Des *Buffers* sont ensuite appliqués sur la précision standardisée, puis dissouts avec l'outil *Dissolve* et son option *Create multipart features* préalablement décochée. De nouvelles prélocalités, sont alors créées et certaines d'entre-elles touchent les polygones du groupement N°10. Une nouvelle sélection par emplacement combinée à un *Summary Statistics* permet de connaître le nombre de ces nouvelles prélocalités issues des notes repêchées qui touchent des polygones du groupement N°10. Parallèlement, une nouvelle sélection par emplacement avec l'option *Switch Selection* sélectionnées suivie à nouveau d'un *Summary Statistics* permet de retrouver le nombre de notes qui n'ont pas pu être « repêchées » même après la standardisation et la fusion de leur *Buffers.* La dernière partie du modèle et très semblable à ce qui a déjà été décrit pour la méthode N°3. Ainsi, l'ensemble des tables obtenues à la suite des *Summary Statistics* sont regroupées en un unique tableau et leurs champs sont renommés de manière à les rendre plus intelligibles. La dernière étape consiste à calculer la fidélité de l'espèce à son groupement selon la formule suivante $\frac{(\text{BUF_INCLUS_GRPMT}+\text{NOTES_REPECHEES})}{\text{NBRE_BUF_TOT}}$.

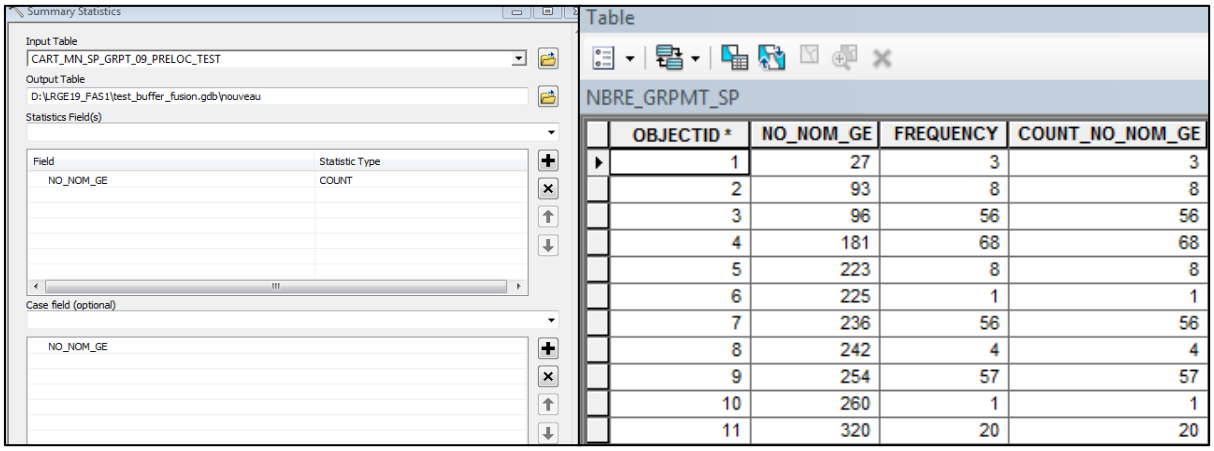
L'exemple ci-dessous (**Figure 26**) présente la table finale obtenue avec *Aceras anthropophorum.*

OBJECTID *	NO_NOM_GE	NBRE_BUF_TOT	BUF_INCLUS_GRPMT	NOTES_REPECHEES	NOTES_HORS_GRPMT	TOTAL_BUFF_GRPMT	CO_GRPMT	FIDELITE_GRPMT
1	11	137	108	7	8	115	10	83.9416

Figure 26 : Exemple de sortie obtenue avec la méthode N°4.

Le NO_NOM_GE renseigne sur l'identité du taxon. Le champ NBRE_BUF_TOT n'est rien d'autre que le décompte du nombre de *Buffers* par espèce obtenus à partir du champ XY_PRECISION des différents taxons. Ce champ correspond aussi au nombre de notes pour chacune des espèces pour l'intervalle de temps [2006-2017]. BUF_INCLUS_GRPMT indique le nombre de notes, par extension de *Buffers*, qui ont été inclus dans leur groupement de milieux naturels respectif, soit celui des prairies et ourlets maigres/secs dans cet exemple. Par ailleurs, un *Buffer* qui ne serait pas totalement inclus au sein du groupement serait aussi comptabilisé. Autrement dit, il ne s'agit pas d'une inclusion stricte. Le champ NOTES_REPECHEES comptabilise les notes qui ont pu être rattachées à leur groupement de milieux naturels après la standardisation de leur précision respective à 25 m suivie de la fusion de leur *Buffer.* Le nombre de notes qui n'ont pas pu être « repêchées » est inscrit sous le champ NOTES_HORS_GRPMT. L'addition du nombre de notes présentes dans les champs BUF_INCLUS_GRPMT et NOTES_REPECHEES permet de renseigner TOTAL_BUFF_GRPMT. Le numéro du groupement est donné sous le champ CO_GRPMT et fera office d'étiquette permettant de tracer l'origine des données lorsque l'ensemble des résultats sera fusionné en un unique tableau de synthèse. La fidélité de l'espèce à son groupement est indiquée au sein de la dernière colonne.

Comme pour la méthode N°3, la création des localités n'est pas effectuée directement au sein du même *Model Builder* mais de façon semi-automatisée à l'aide du modèle 9_MB_SP_MN_PRELOC. Même si cela n'a pas été fait au sien du modèle, il est tout à fait envisageable de connaître le nombre de polygones de milieux naturels pour chacune des espèces à l'aide d'un *Summary Statistics* appliqué sur le champ NO_NOM_GE comme démontré dans la **Figure 27** ci-dessous.



OBJECTID *	NO_NOM_GE	FREQUENCY	COUNT_NO_NOM_GE
1	27	3	3
2	93	8	8
3	96	56	56
4	181	68	68
5	223	8	8
6	225	1	1
7	236	56	56
8	242	4	4
9	254	57	57
10	260	1	1
11	320	20	20

Figure 27 : Méthode « manuelle » pour connaître le nombre de prélocalités par taxon et par groupement de milieux naturels. Un extrait du résultat est donné dans le tableau de droite.

3.9 Création semi-automatisée des localités

A partir des sorties des modèles des méthodes N°3 et N°4, il est possible d'obtenir des cartes de localités de façon semi-automatisée à l'aide du *Model Builder* 9_MB_SP_MN_PRELOC présenté à la page 32. La couche en entrée du modèle est obtenue suite aux étapes de la jointure spatiale combinée à la suppression des doublons **(Figure 28)** via les différents modèles des méthodes N°3 et N°4.

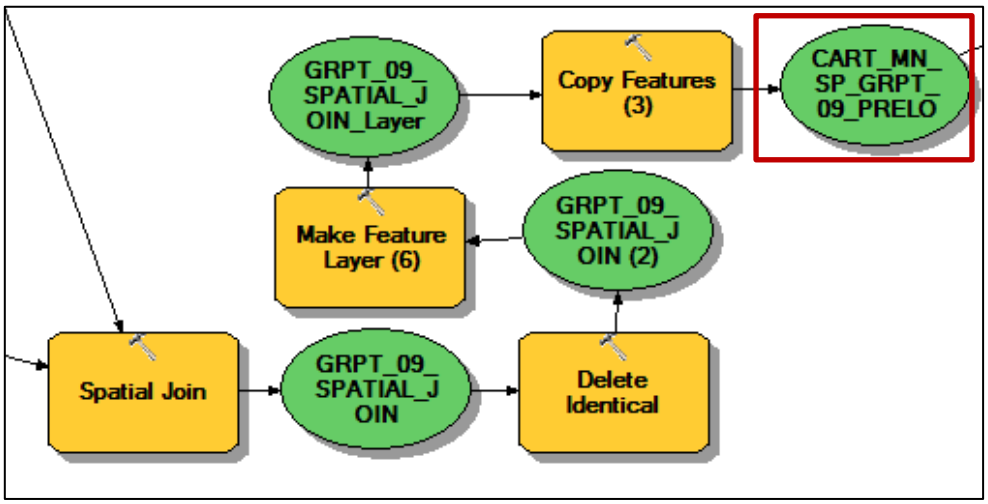


Figure 28 : L'encadré rouge représente le fichier de sortie à introduire dans le modèle 9_MB_SP_MN_PRELOC pour l'obtention des prélocalités pour chacun des taxons de façon automatique.

Les couches **CART_MN_SP_GRPT_XX_PRELOC** obtenues pour l'ensemble des 11 modèles des deux méthodes contiennent l'information sur la composition taxonomique de chacun des polygones de groupements de milieux naturels après la suppression des éventuels doublons. L'exemple ci-dessous présente la composition floristique d'un polygone du groupement N°9. Pour rendre l'exemple plus lisible, seul les champs utiles à la compréhension de cet exemple ont été représentés (**Figure 29**).

Table

CART_MN_SP_GRPT_09_PRELOC

OBJECTID *	SHAPE *	Join_Count	TARGET_FID	JOIN_FID	NO_NOM_IF	NOM_COMPLET	NO_NOM_GE
1	Polygon	1	7	42	1187	Hippocrepis emerus (L.) Lassen	1213
2	Polygon	1	7	53	2779	Carex digitata L.	451
3	Polygon	1	7	54	2784	Carex alba Scop.	442
4	Polygon	1	7	58	144	Thalictrum aquilegiifolium L.	2454
5	Polygon	1	7	59	69	Dryopteris carthusiana (Vill.) H. P. Fuchs	784
6	Polygon	1	7	60	1372	Mercurialis perennis L.	1529
7	Polygon	1	7	61	3122	Paris quadrifolia L.	1707
8	Polygon	1	7	62	66	Dryopteris filix-mas (L.) Schott	787

Figure 29 : Exemple du contenu d'une table introduite en entrée du modèle 9_MB_SP_MN_PRELOC. Le champ TARGET_FID corresponds à l'identifiant du polygone de groupements MN.

Le polygone 7 est ainsi composé de 8 espèces regroupées dans le groupement N°9 soit celui des forêts et ourlets mésophiles. A nouveau, le nom complet permet de s'assurer qu'il n'y a pas eu d'éventuelles coquilles ce qui ne semble pas être le cas ici puisque l'ensemble des 8 espèces composant le polygone 7 sont toutes des espèces calcicoles forestières ou de lisières. Cependant, connaître l'ensemble des espèces composant un polygone de milieu naturel ne nous intéresse pas dans l'immédiat. Nous souhaiterions plutôt connaître les différentes localités de *Carex alba* Scop. ou de *Carex digitata* L. présentent sur l'ensemble du canton et générer une carte de localités par espèce. Cela est possible en introduisant l'ensemble de la couche à l'intérieur du modèle 9_MB_SP_MN_PRELOC présenté dans la **Figure 30**

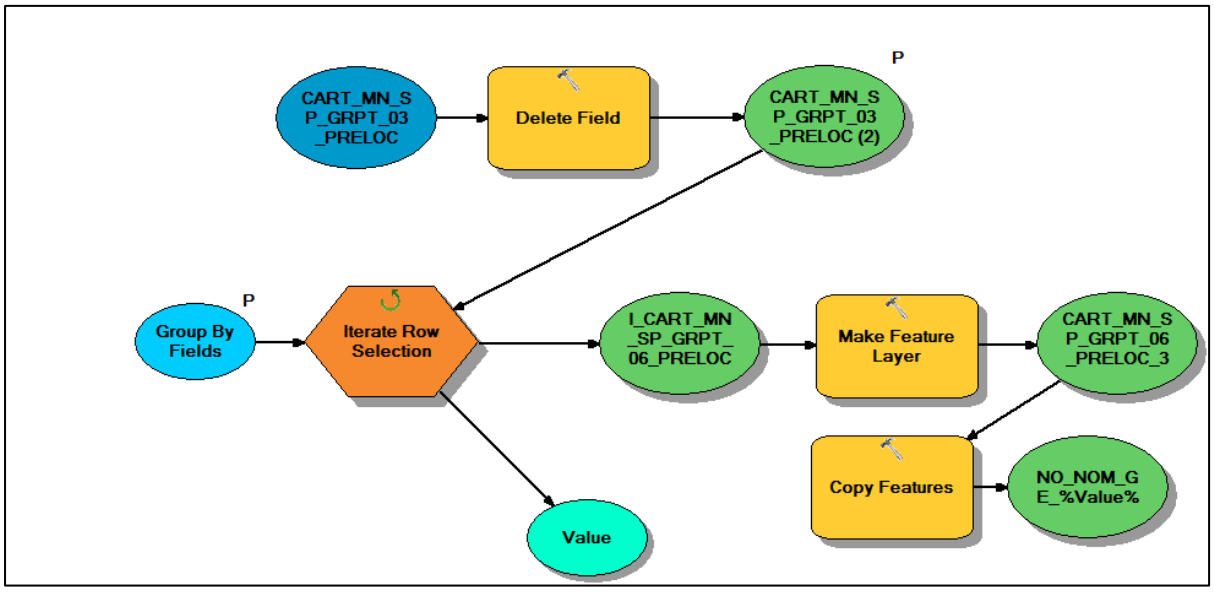


Figure 30 : *ModelBuilder* du modèle 9_MB_SP_MN_PRELOC permettant la génération automatisée des prélocalités.

Contrairement aux autres *Model Builder* construit précédemment, il suffit de changer la couche en entrée pour refaire tourner le modèle. Après avoir éliminé les champs inutiles via l'outil *Delete Field*, un *Iterate Row Selection* basé sur le champ NO_NOM_GE permet de créer des cartes de polygones pour chacun des taxons du groupement de milieux naturels introduits en entrée. Les cartes obtenues seront nommées par l'identifiant genevois du taxon. La façon d'écrire le cheminement du dossier de destination, D:\LRGE19_FAS1\Test.gdb\NO_NOM_GE_%Value%, est ici crucial. Le signe pourcent permet d'éviter que les couches obtenues à chaque « run » de l'itérateur n'écrasent les précédentes lors de leur création. Les différentes couches générées sont ainsi ajoutées les unes après les autres au sein d'une *geodatabase* présélectionnées. Il est ensuite possible d'effectuer les mêmes opérations avec

les couches de *Buffers.* Finalement, pour obtenir une carte de localités d'un taxon, il suffit d'afficher sa couche de polygones et sa couche de *Buffers* de manière simultanée et de tracer à la main les localités sur la base de connaissances acquises par la pratique et sur le terrain.

4. Résultats obtenus via *ModelBuilder*

Comme expliqué en plus haut, il n'existe pas de méthodes uniques et réellement efficaces pour la création des localités parce que la définition même de ce terme est soumise à interprétation. Le concept de menace commune, sur lequel repose en partie la définition de localité, est également difficile à appréhender par la modélisation. Bien que la méthode N° 4 soit la plus aboutie, les trois autres peuvent aussi apporter quelques avantages. Dans tous les cas, l'utilisateur des différents modèles n'échappera pas à regrouper les localités de façon plus au moins empirique. Les cartes ci-dessous représentent des exemples de résultats obtenus à l'aide des modèles des différentes méthodes.

4.1 Exemples de résultats de créations de prélocalités obtenus avec les 4 méthodes

Les pages suivantes ont pour but de présenter les résultats (**Figures 32-35**) obtenus avec les différentes méthodes dans la création des prélocalités d'*Acers anthropophorum*, une orchidée relativement commune dans le canton de Genève mais néanmoins présente dans moins de 60 mailles 500 **(Figure 31).** Le terme prélocalité est à mon avis plus juste dans le sens où il n'y a pas de traitement géomatique automatisé pour créer et décompter les localités.

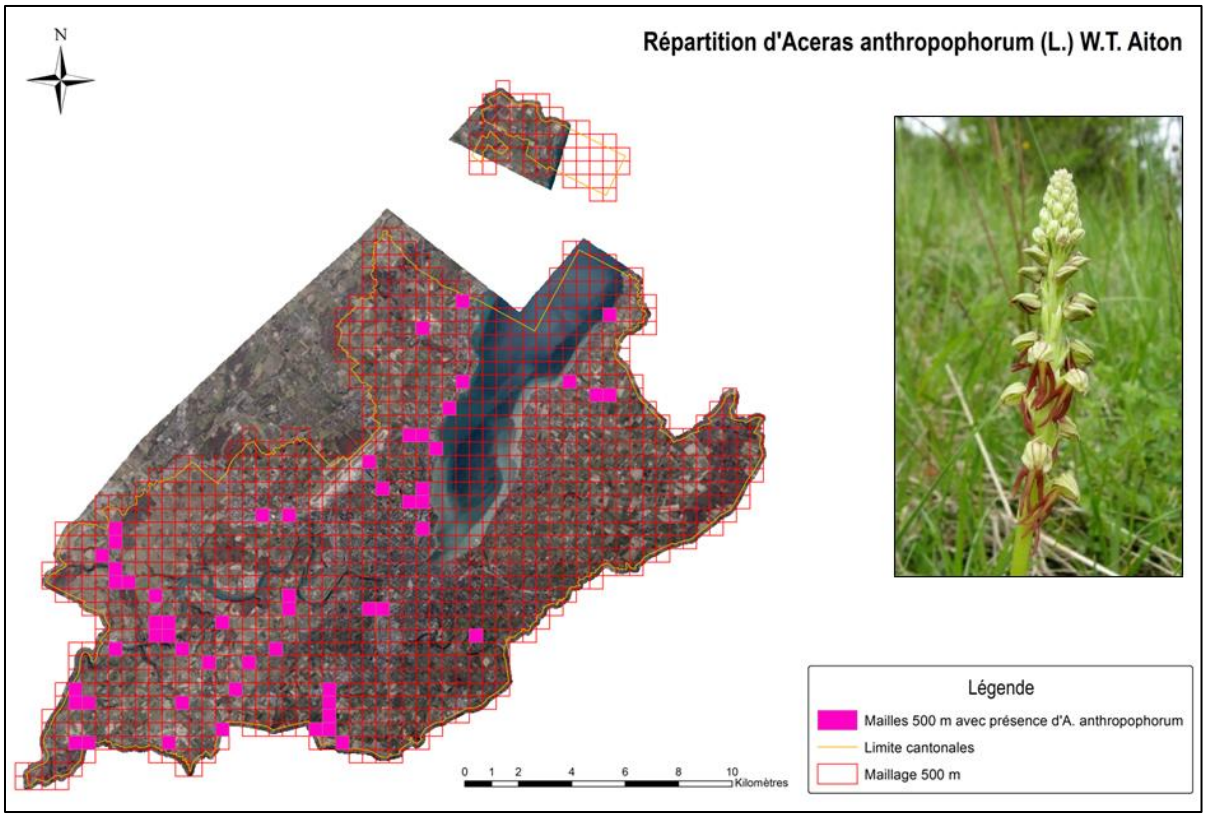


Figure 31 : Répartition d'*Aceras anthropophorum* dans le canton de Genève pour l'intervalle de temps [2006-2017]. L'espèce a subi d'une légère diminution dans sa fréquence entre les deux pas de temps et elle est présente dans moins de 15 km^2.

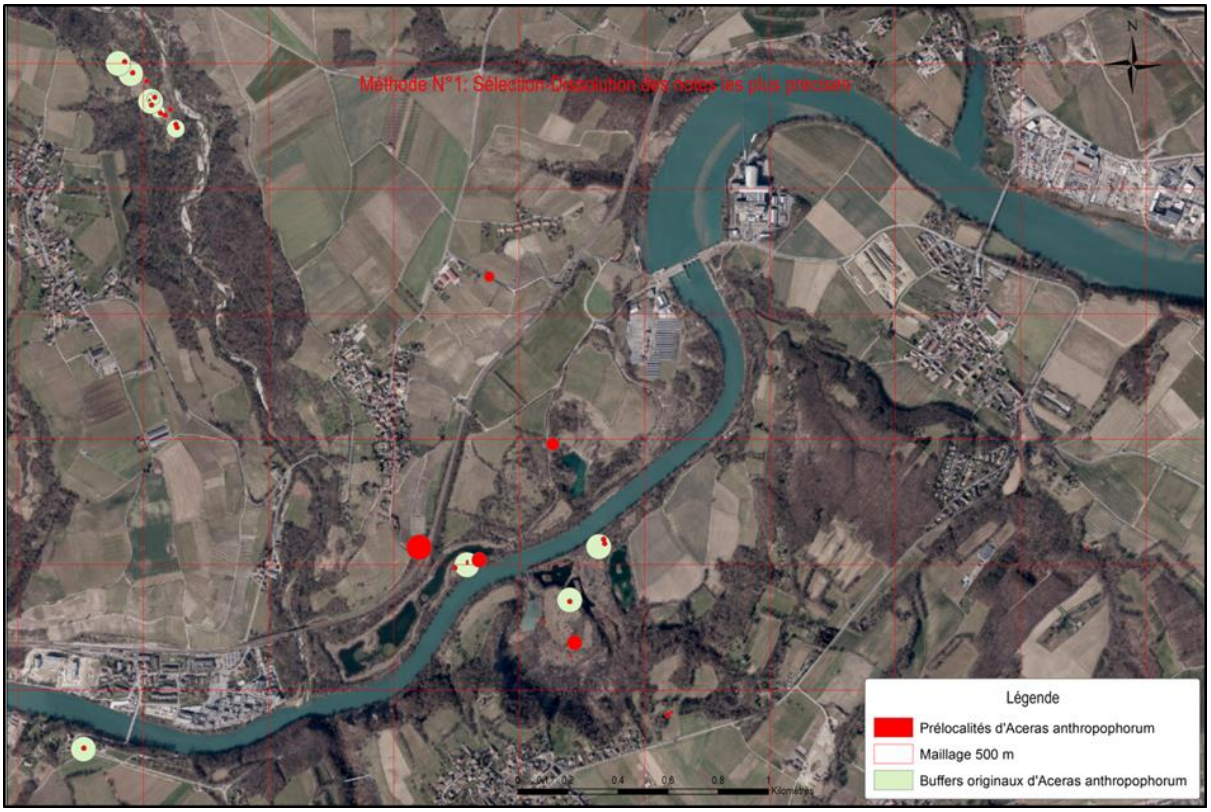


Figure 32 : Carte des prélocalités d'*A. anthropophorum* obtenues avec la méthode N°1 dans le Vallon de l'Allondon et la région de Cartigny.

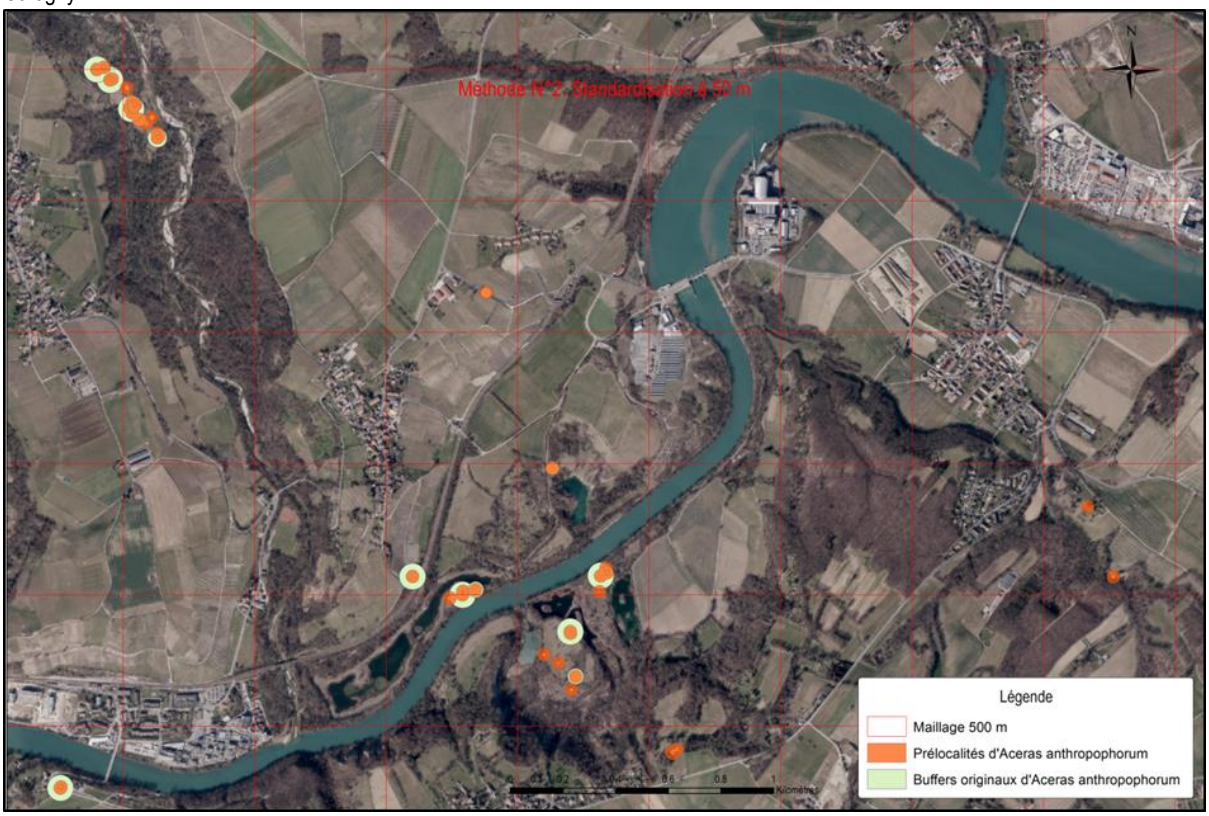


Figure 33 : Carte des prélocalités d'*A. anthropophorum* obtenues avec la méthode N°2 dans le Vallon de l'Allondon et la région de Cartigny.

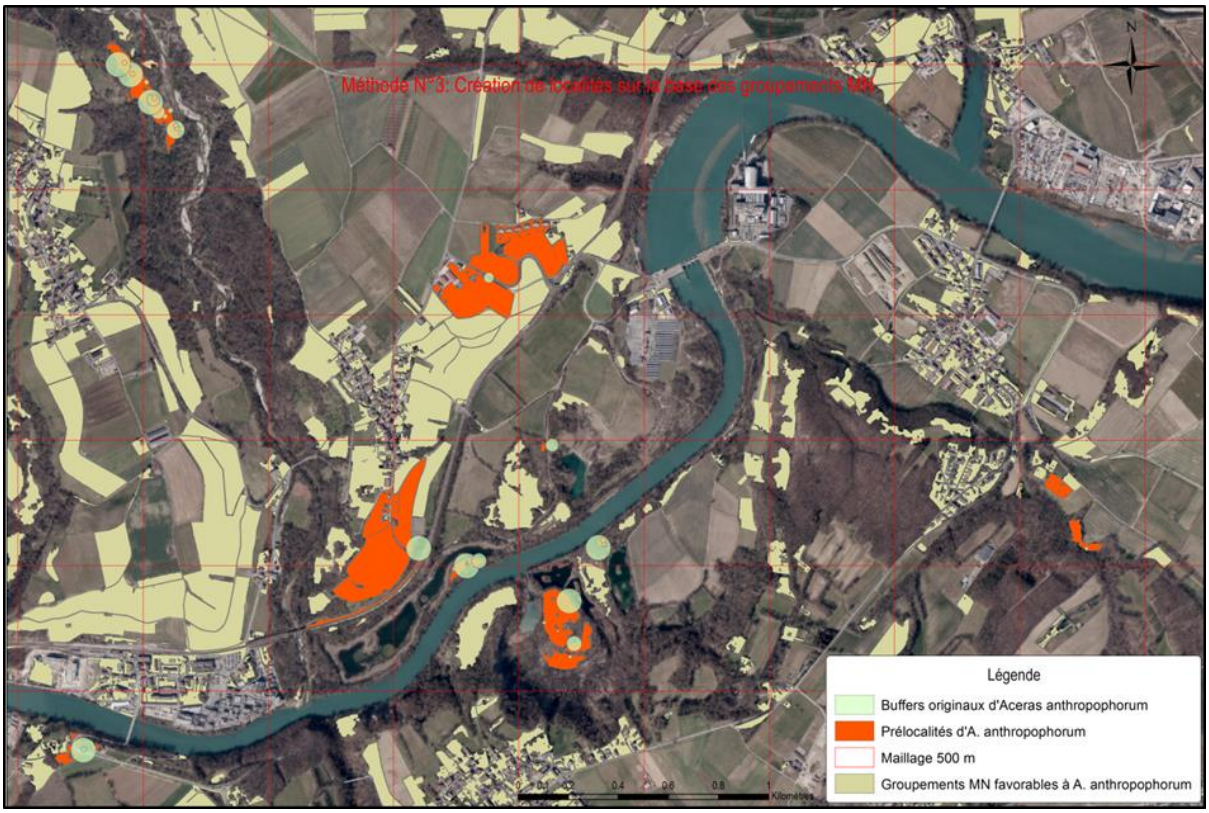


Figure 34 : Carte des prélocalités d'*A. anthropophorum* obtenues avec la méthode N°3 dans le Vallon de l'Allondon et la région de Cartigny.

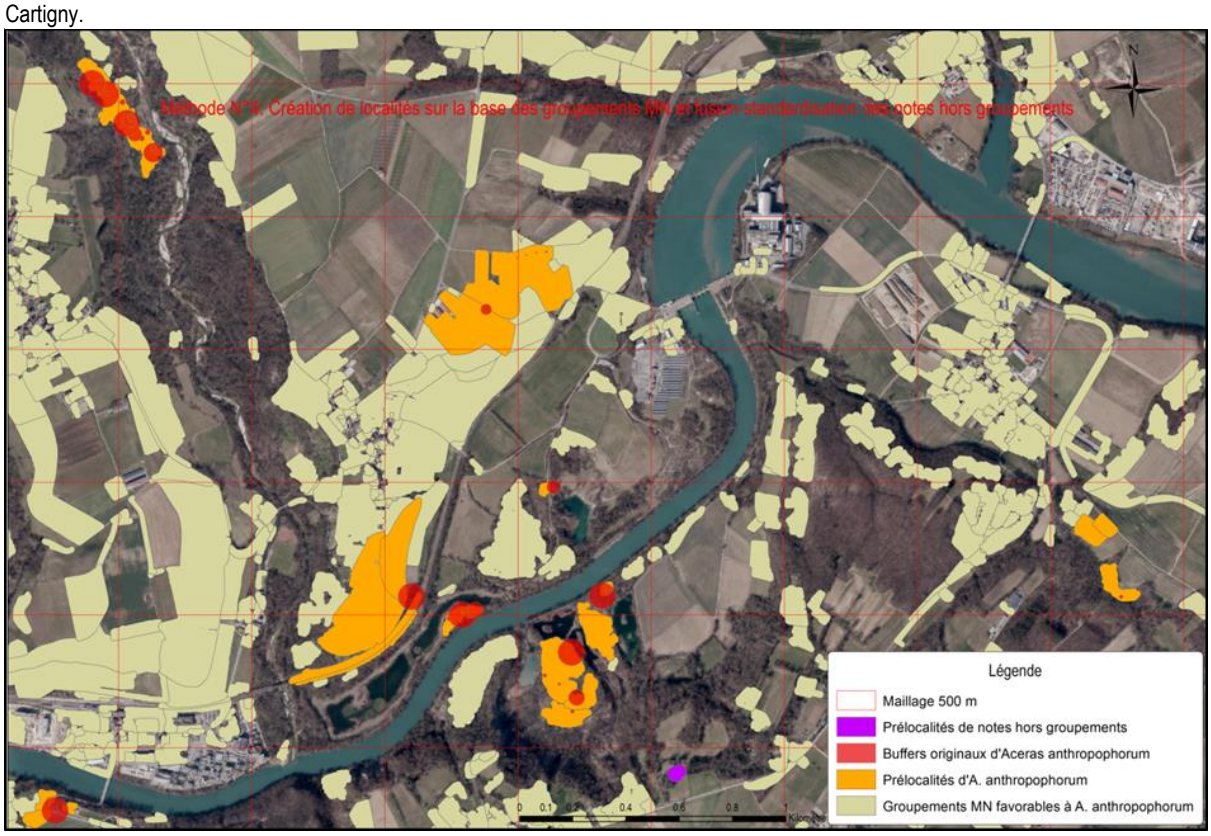


Figure 35 : Carte des prélocalités d'*A. anthropophorum* obtenues avec la méthode N°4 dans le Vallon de l'Allondon et la région de Cartigny.

D'un point de vue strictement visuel, la méthode N°4 est probablement la plus informative. Les prélocalités apparaissent nettement grâce à l'utilisation des groupements de milieux naturels et il devient aisé de les regrouper pour en faire des localités à condition bien évidemment d'avoir une connaissance préalable du terrain.

4.2 Exemple concret de création de localités avec la méthode N°4

La carte ci-dessous présente (**Figure 36**) un exemple de regroupement des prélocalités d'*Aceras anthropophorum* en 10 localités. Le fait de rajouter certaines couches supplémentaires sur la carte comme les limites de réserves naturelles, **SDO.AGGLO_NAT_RESERVES**, ou bien encore les zones alluviales, **SDO.AGGLO_NAT_ZONES_ALLUVIALES**, donne des informations complémentaires bien utiles à la délimitation des différentes localités car certaines menaces peuvent en être extrapolées.

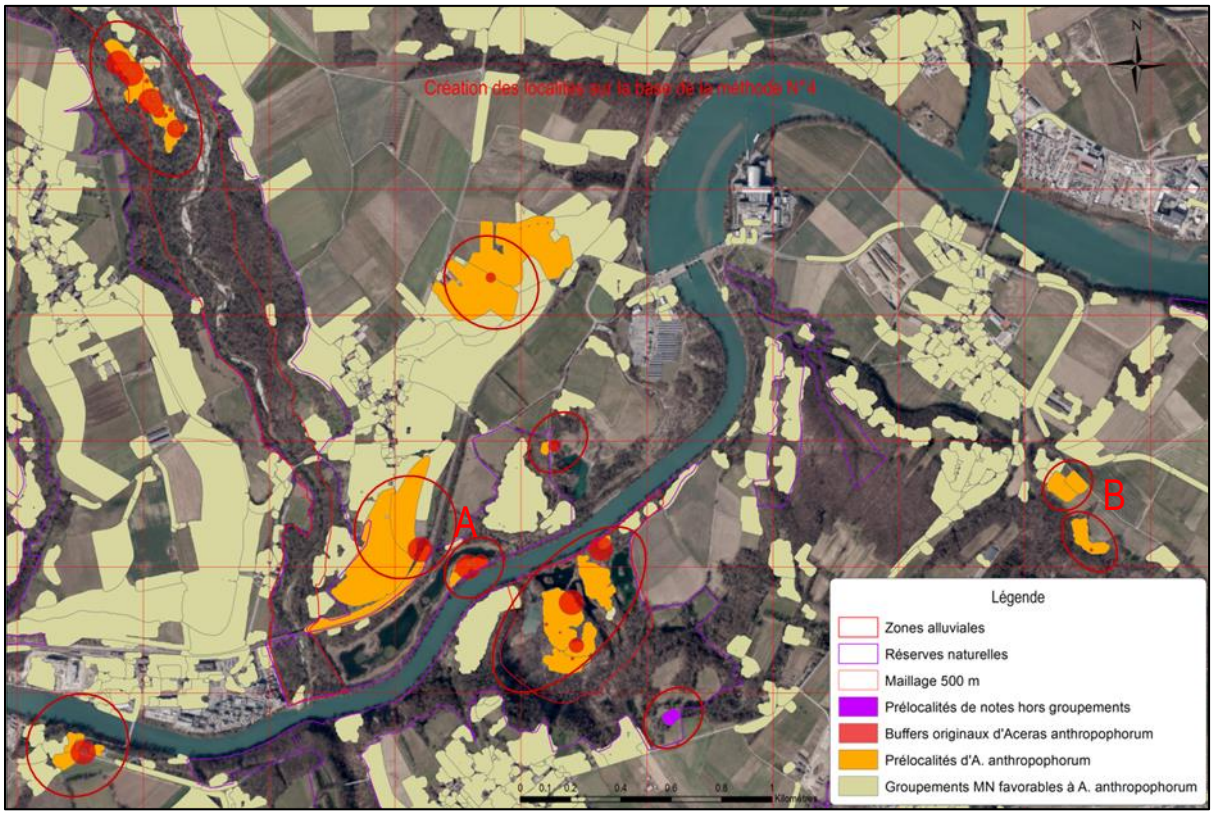


Figure 36 : Cartes des localités d'*Aceras anthropophorum* obtenues via la méthode N°4 dans le Vallon de l'Allondon et la région de Cartigny. L'ajout des couches Réserves naturelles et Zones alluviales permettent de délimiter plus facilement les localités sur une base visuelle.

Dans l'exemple ci-dessus, l'importance de l'utilisation des couches **SDO.AGGLO_NAT_RESERVES** et **SDO.AGGLO_NAT_ZONES_ALLUVIALES** dans la délimitation des localités est démontrée. Certaines prélocalités ne sont distantes que d'une centaine de mètres mais le fait que l'une se trouve au sein d'une réserve naturel et l'autre sur un talus de chemin de fer [A] pourrait potentiellement indiquer une gestion différente des stations où l'espèce croît ce qui pourrait justifier la création de deux localités. Pour certaines stations spatialement proches [B] le regroupement des prélocalités en une seule localité pourrait être sujet à d'avantage de discussions. Idéalement, il faudrait connaître le nombre de propriétaires des différentes parcelles et savoir si leur mode de gestion est semblable ce qui pourrait justifier le regroupement des prélocalités en une unique localité. Dans le cas où les prélocalités seraient considérées comme deux entités indépendantes, le plus grand risque consisterait en une augmentation

artificielle du nombre de localités du taxon qui engendrerait une sous-estimation de son statut de menace.

5. Problèmes rencontrés avec *ModelBuilder* : quelques recommandations

Dans cette partie, je souhaiterais énoncer brièvement les différents problèmes que j'ai rencontrés lors de la conception des *ModelBuilder* et les moyens mis en œuvre pour les contourner.

Les sélections par attributs (*Select Layer By Attribute*) partielles

Le manque de stabilité des différents modèles a été un problème récurrent tout au long de leur conception. Ce problème de stabilité est particulièrement patent lorsque le modèle est amené à effectuer une dichotomie entre deux intervalles de temps à l'aide d'une sélection par attribut. En effet, il arrive fréquemment que le modèle ne « fasse le travail » que pour un seul des deux intervalles de temps. L'anomalie peut être corrigée, du moins en partie, en effectuant une copie de la sélection par attribut. Le fait de créer des copies de chacune des étapes permet aussi de ne pas avoir à refaire tourner l'ensemble du modèle en repartant directement de celles-ci après avoir déconnecté les parties du modèle plus en amont. De la même façon, la sélection par attribut dysfonctionne de façon quasi récurrente lors de la sélection de NULL. Il faut alors faire tourner une seconde, voire une troisième fois le modèle pour que la sélection par attribut des NULL soit enfin effectuée. Malheureusement, il faut aussi supprimer l'ensemble des fichiers déjà créés par le modèle et cocher-décocher l'ensemble des options des différents outils avant de pouvoir refaire à nouveau tourner le modèle sous peine de le bloquer à certaines étapes ou pis, de se retrouver avec des tables finales dans lesquelles les informations ont été dupliquées (voire tripliquées, quadruplées…). Ces problèmes rencontrés avec les sélections par attribut sont d'autant plus frustrants qu'ils n'interviennent pas lorsque les sélections sont effectuées manuellement, autrement dit en dehors du *ModelBuilder*. Mais faire l'ensemble des opérations en dehors du *ModelBuilder* reviendrait à renoncer à toutes automatisations.

La standardisation des zones tampons (*Buffers*) impossible

Lors de la standardisation de la précision des notes à 25 m et de la création des *Buffers* qui en découle, il arrive fréquemment que les nouvelles zones tampons soient produites à partir du champ XY_PRECISION et non du nouveau champ PRECISION. Pour contrer ce problème, il faut introduire en amont de l'outil *Buffer* un *Delete Field* pour supprimer l'ensemble des champs qui se rapportent à l'ancienne précision. Si le modèle a déjà tourné une fois avant cette correction, le résultat obtenu n'est pas garanti. Il faut alors supprimer les champs qui se rapportent à l'ancienne précision de façon manuelle, puis réintroduire le tableau ainsi obtenu dans le modèle en ayant pris soins de cocher-décocher l'ensemble des options des différents outils et de renommer voire supprimer dans la *geodatabase* les couches et les tableaux déjà créés.

Eviter les copies inutiles

Effectuer une copie directement après l'utilisation de l'outil *Buffer* n'est pas indiqué sous peine de se retrouver avec un modèle qui tourne en boucle. L'outil *Buffer* créant déjà une nouvelle couche, les données introduites dans l'outil ne sont pas modifiées. Dès lors, effectuer une copie de la couche des *Buffers* n'a pas beaucoup de sens. Cependant, j'ai aussi remarqué que lorsque le modèle tournait une

première fois, la copie de la couche des *Buffers* s'effectuait sans problème. A nouveau, les problèmes surviennent lorsque les modèles sont *runnés* plusieurs fois de suite.

Le manque de souplesse du *Model Builder* face à la répétition

Hormis pour le modèle 9_MB_SP_MN_PRELOC, l'utilisateur n'aura pas d'autres choix que de cocher-décocher l'ensemble des options des différents outils contenus au sein des modèles avant de les refaire tourner s'il veut s'éviter les problèmes énoncés ci-dessus. Il en résulte une perte de temps relativement conséquente et une incertitude quant à la réussite des différentes sélections par attributs.

Eviter les modèles trop grands et trop complexes

Il est préférable de construire un ensemble de petits modèles plutôt que d'en créer un grand qui ne fonctionne que partiellement et dans lequel les erreurs potentielles sont difficilement repérables. Malheureusement, le « saucissonnage » nuit à l'automatisation.

6. Passage sur FME

A cause du manque de stabilité de certaines parties du modèle, il a été décidé de passer par la suite sur FME (*Feature Manipulation Engine*). FME est un logiciel de traitement de la famille des ETL (*Extract Transform Load*). Il permet la manipulation de données spatiales mais aussi d'images. Les ETL sont souvent utilisés pour effectuer des synchronisations massives d'information d'une source de données, par exemple une base de données, vers une autre. Dans FME, de nombreux *transformers* permettent de manipuler les données (agrégations, filtres, conversions, suppressions…) entre le réservoir source (*Reader*), qui peut être une *geodatabase*, et le réservoir cible (*Writer*). Les quelques lignes et figures suivantes ont pour but de décrire brièvement les quelques tests que j'ai effectués sur FME. Malheureusement, je n'ai pas réussi à transposer l'ensemble des modèles créés sur ArcGIS sur FME par manque de temps. Les opérations présentées ci-dessous sont donc basiques.

6.1 Attribution des statuts de menace provisoires selon le critère A avec FME

Le modèle conçu sur FME pour l'attribution des statuts de menaces provisoires est très semblable au *Model Builder* créé sur ArcGIS. Les données utilisées comme *Reader* (données sources) proviennent d'une table de synthèse regroupant le nombre d'occurrences, autrement dit le nombre de mailles kilométriques, pour les intervalles de temps [1990-2005] et [2006-2017]. La table contient aussi les occurrences calculées sur la base du maillage de 500 m pour le deuxième intervalle de temps.

Le modèle (**Figure 37**) repose sur une suite de sélections suivies des attributions des statuts de menace.

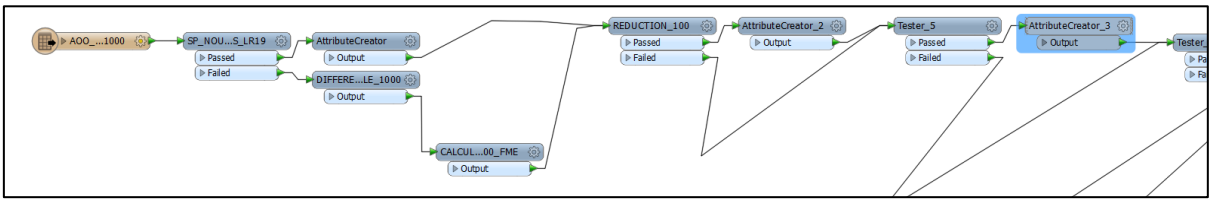


Figure37 : Détail sur le modèle FME utilisé pour l'attribution des statuts de menace provisoires.

La première étape consiste à sélectionner les taxons pour lesquels il n'y avait pas de données pour le premier intervalle de temps, autrement dit les espèces nouvelles pour notre flore. L'opération est possible à l'aide du *transformer Tester* qui est le pendant du *Select Layer By Attribute* d'ESRI. Dans

cette première étape, les taxons qui n'existent pas dans le premier intervalle de temps, autrement dit dont le champ TAB_1000_90_05_1000_AOO1000_OLD est renseigné par un zéro, « passent le test » alors que les autres « échouent » et se retrouvent dirigés vers un *ExpressionEvaluator*. Les données qui « ont passé le test » sont alors dirigées vers un *AttributeCreator* qui permet de leur assigner leur statut de menace, un zéro dans le cas présent parce que l'évolution des occurrences de ces taxons ne peut pas être calculée. Ce *transformer* permets aussi d'ajouter le champ STATUT_PROV_A au sein du tableau de synthèse. Le *transformer ExpressionEvaluator* (**Figure 38**) permet d'ajouter un nouveau champ au tableau de synthèse, DELTA_NOMBRE _DE_MAILLES_FME, puis de le renseigner directement par les résultats de la soustraction entre les occurrences de chacun des taxons pour les deux intervalles de temps.

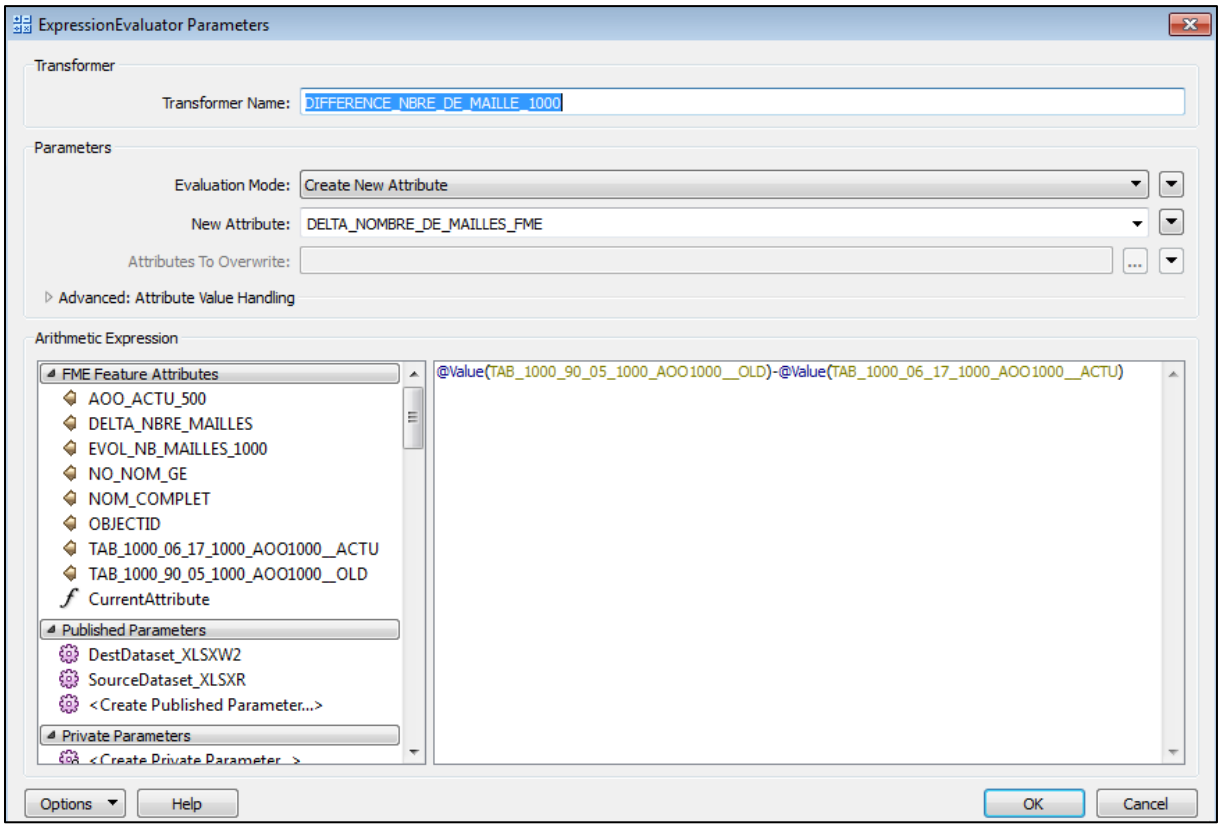


Figure 38 : Le *transformer ExpressionEvaluator* permet d'ajouter un nouveau champ, puis de le renseigner immédiatement.

Une fois le champ DELTA_NOMBRE_DE_MAILLES_FME renseigné, un nouveau champ, CALCUL_EVOLUTION_MAILLES_1000_FME est ajouté à la table de synthèse (**Figure 39**). Ce nouveau champ permet de calculer, en pourcent, l'évolution du nombre d'occurrences de chacune des espèces entre les deux pas de temps de manière identique à ce qui a été fait au sein du *ModelBuilder*. La première dichotomie effectuée tout au début du modèle prend ici tout son sens ! Une division par zéro étant impossible dans les Réels, il fallait écarter le plus en amont possible les taxons pour lesquels il n'y avait pas de données pour le premier intervalle de temps. Les autres étapes du modèle consistent simplement à donner un statut de menace à chacun des taxons en fonction de l'évolution de leur nombre d'occurrences entre les deux intervalles de temps à l'aide de la combinaison des *transformers Tester* et *AttributeCreator* mis bout à bout.

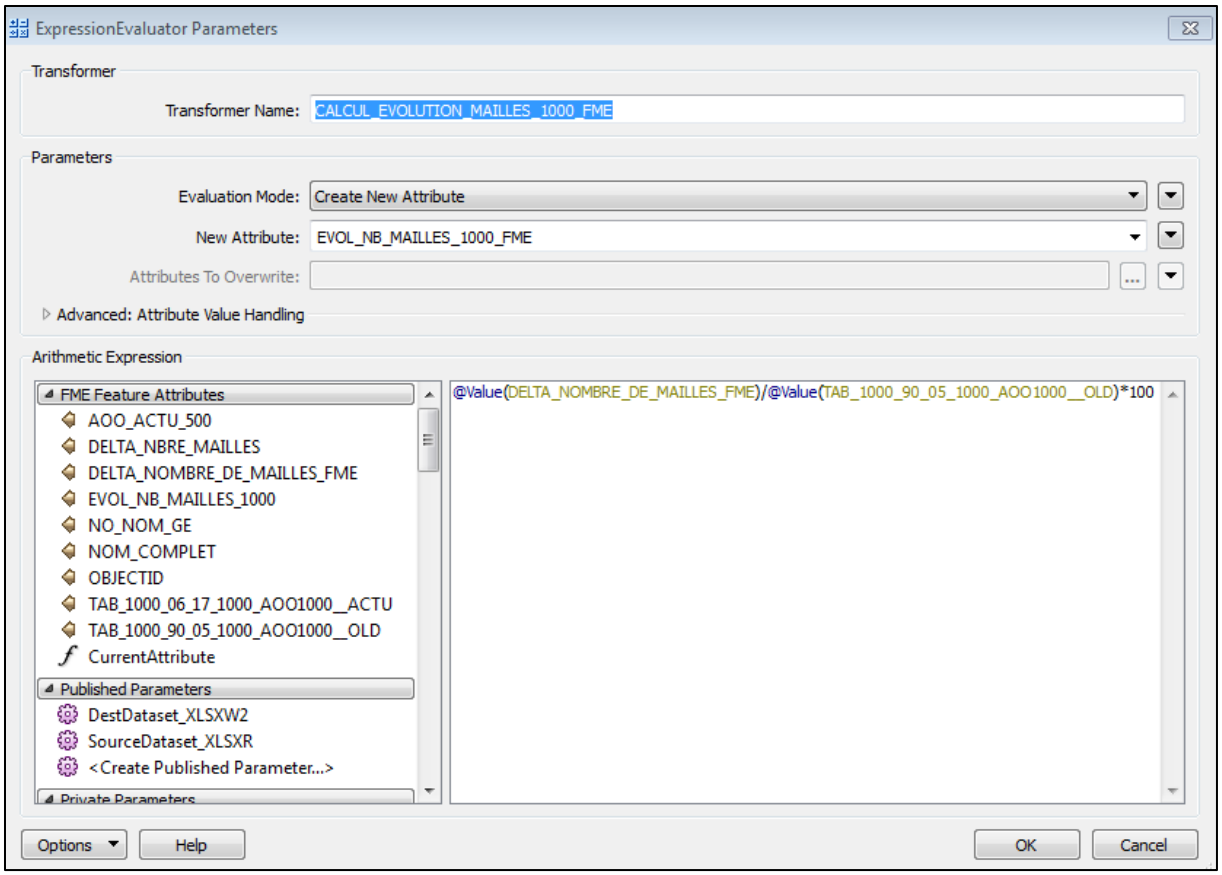


Figure 39 : L'évolution du nombre de mailles est calculée à l'aide d'un *ExpressionEvaluator.*

La figure ci-dessous (**Figure 40**) présente la façon de remplir un *Tester* lorsqu'il s'agit d'évaluer un intervalle.

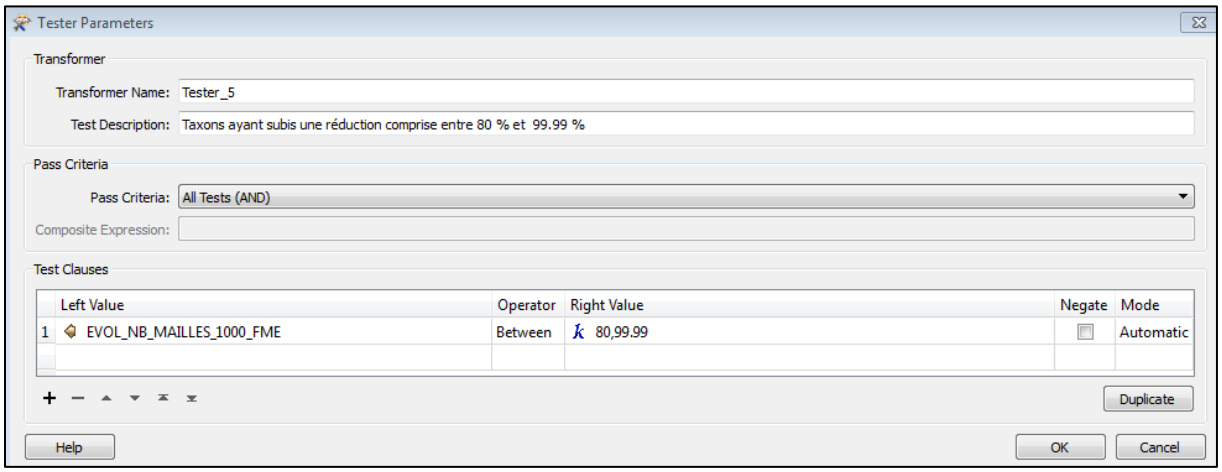


Figure 40 : Exemple d'une sélection sur un intervalle [80 ; 99.99] à l'aide d'un *Tester*.

L'ensemble des résultats est finalement retranscrit au sein d'un fichier Excel, qui fait ici office de *Writer,* directement sur le Bureau et directement disponible pour d'éventuelles analyses.

6.2 Préparation des données sur FME

Il peut paraître surprenant de retrouver la préparation des données à la suite de la section décrivant l'attribution des statuts de menace aux différents taxons, mais n'ayant jamais utilisé FME par le passé, il

me fallait commencer par des traitements de données ultrabasiques avant de me lancer dans d'autres à peine plus compliqués.

La première étape du modèle (**Figure 41**) a pour but de sélectionner uniquement les trachéophytes de la table **CJB.SIPV_FS_OBS** qui fait office ici de *Reader* après avoir éliminé les champs inutiles pour la suite des traitements géomatiques à l'aide du *transformer AttributeKeeper*. Les autres sélections, toutes effectuées à l'aide du *transformer Tester*, sont identiques à celles faites au sein du premier Model Builder. Par conséquent, il ne me semble pas très utile de les développer à nouveau dans cette section.

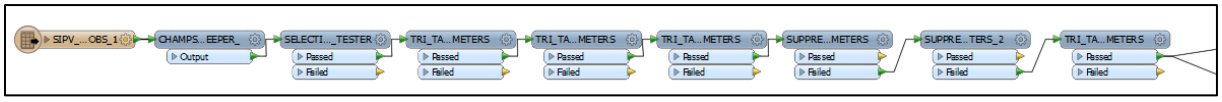


Figure 41 : Détail sur la première partie du modèle FME utilisé pour la génération des données de base.

La seconde étape (**Figure 42**) permet de filtrer les données qui ne sont présentes qu'au sein de la dition à l'aide du *transformer SpatialFilter* (**Figure 43a)** qui utilise le maillage kilométrique à la manière d'un tamis.

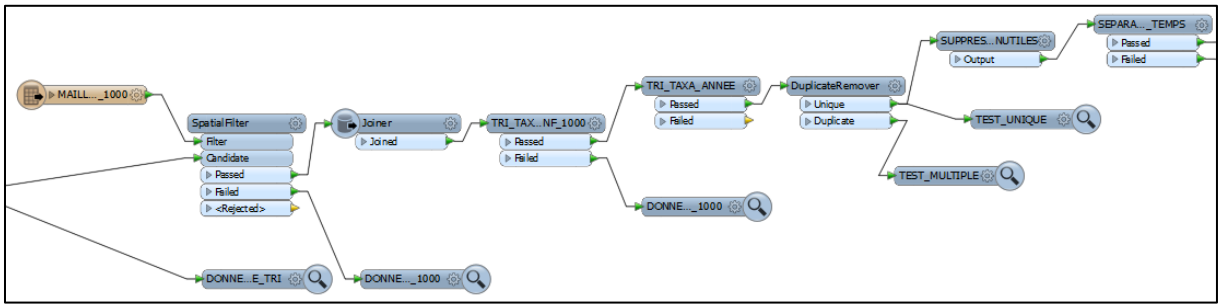
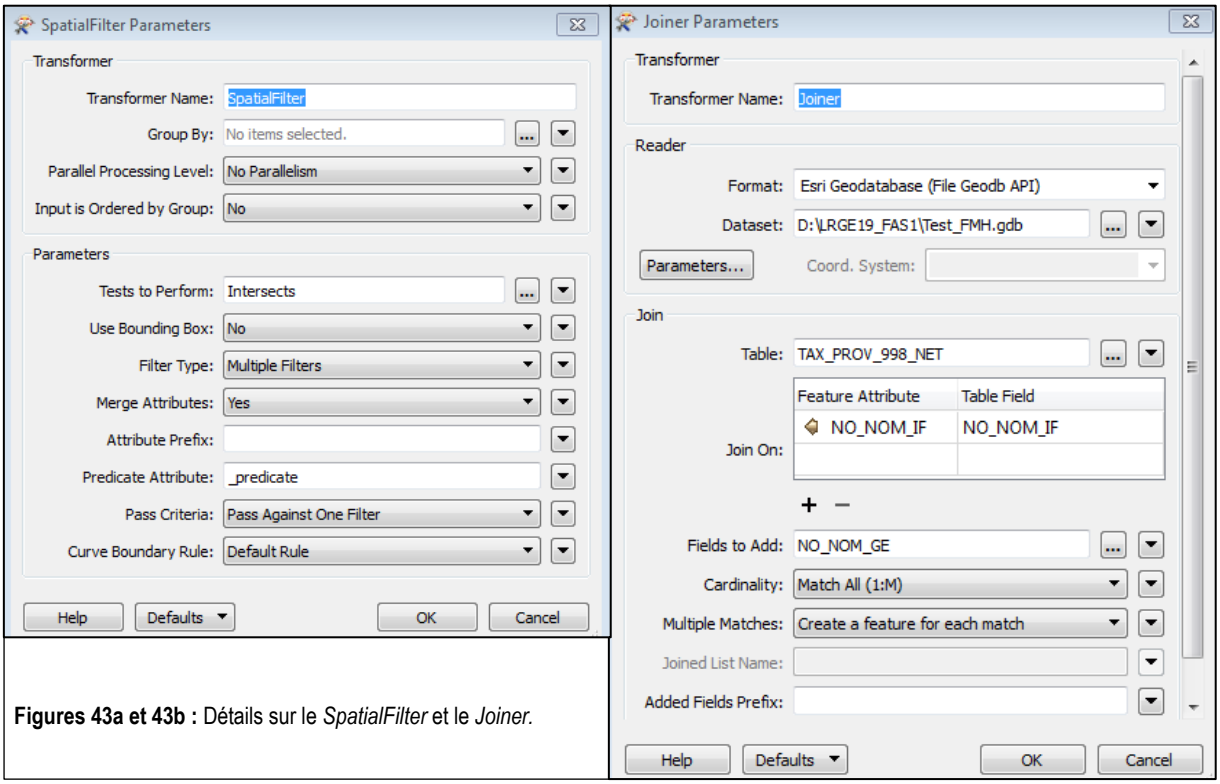


Figure 42 : Suite du modèle FME utilisé pour la génération des données de base.

Le *transformer Joiner* (**Figure 43b**) permet de récupérer les taxons qui seront inclus dans la nouvelle liste rouge ainsi que leur identifiant genevois en joignant la table **CJB.SIPV_FS_OBS** à la table **TAX_PROV_998_NET**. Puis deux sélections sont effectuées pour ne retenir que les données dont la précision est égale ou inférieure à 1000 m et prises après 1989. Certaines données pourraient se trouver à double. Afin de s'en débarrasser, un *DuplicateRemover* basé sur les champs A_NOTE, NO_NOM_GE, NO_NOM_IF, NO_NOTE, X, Y et XY_PRECISION. A nouveau, les champs inutiles rajoutés lors du joint entre les deux tables sont éliminés à l'aide de l'*AttributeKeeper*. Un *Tester* permet ensuite d'effectuer une dichotomie basée sur l'intervalle de temps.



Figures 43a et 43b : Détails sur le *SpatialFilter* et le *Joiner.*

La troisième étape (**Figure 44**) permet d'attribuer à l'aide du *transformer StatisticsCalculator* une fréquence pour chacun des taxons pour les deux types de maillages et les deux intervalles de temps après avoir préalablement supprimé les espèces qui apparaissent plusieurs fois au sein d'une même maille.

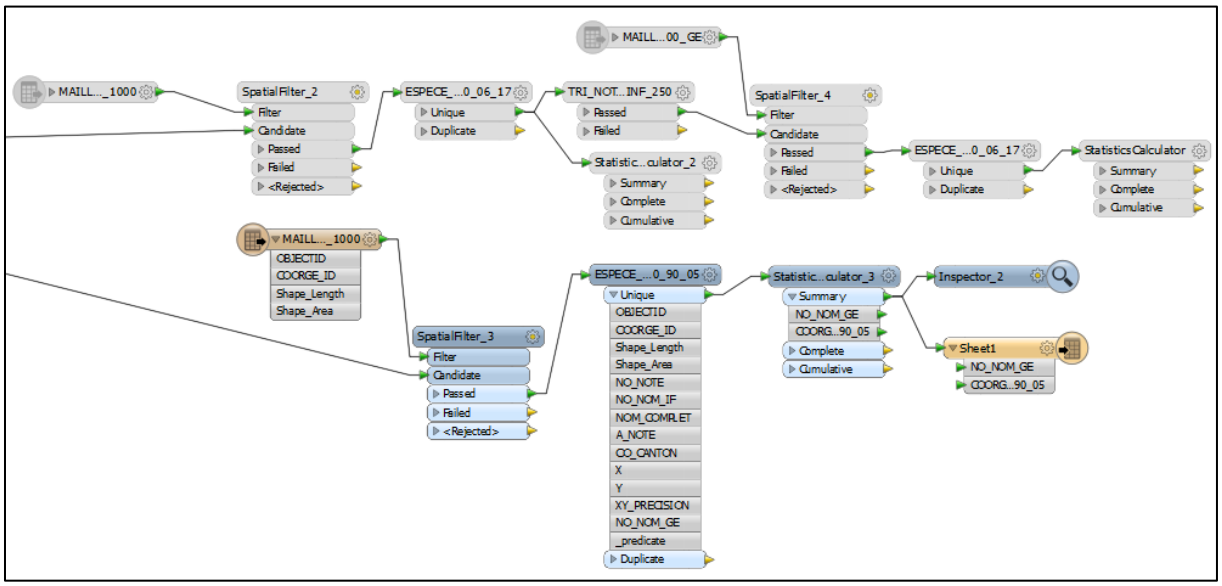


Figure 44 : Détail sur la troisième et dernière partie du modèle FME utilisé pour la génération des données de base.

Idéalement, il faudrait encore réunir l'ensemble des tables obtenues au sein d'un tableau de synthèse ce qui permettrait ensuite l'attribution des statuts de menace provisoires.

7. Discussion

Avantages de FME pour la préparation des données

Les modèles créés sur FME possèdent une souplesse d'utilisation par rapport à ceux construits via *ModelBuilder.* En effet, les modèles peuvent être *runnés* plusieurs fois de suite sans avoir à réactiver l'ensemble des *transformers* comme cela serait le cas pour les options des outils composant les *ModelBuilder*. L'utilisation de préconditions est également inutile sur FME et les opérations sont effectuées de façon logique, dans l'ordre de leur apparition. Les dichotomies sont toujours effectuées correctement. Il est aussi possible de ne faire tourner qu'une partie du modèle et de rajouter à loisir des *transformers* par la suite sans avoir à craindre un *bug*. Le résultat de chacune des étapes apparaît immédiatement, sans avoir à créer de nouveau tableaux comme nous le ferions avec un *Summary Statistics* au sein d'un *ModelBuilder*. Dans le même ordre d'idées, FME ne génère pas une myriade de fichiers en sortie contrairement au produit d'ESRI où il arrive qu'il devienne rapidement difficile de se retrouver dans les différentes *geodatabases*. Ce « problème » est encore plus patent lors des phases d'essais où une multitude de fichiers sont générés. Dans le cas où il serait souhaitable d'avoir d'avantage d'informations, l'utilisation d'un *Inspector* permet d'en savoir un peu plus sur ce qui sort d'un *transformer*.

FME semble aussi plus indiqué pour traiter de grands jeux de données. Avec 4.9 millions d'entrées en 2015 contre 90 000 en 2001 (**Bornand et al., 2016**), les données disponibles sur Info Flora sont de plus en plus nombreuses. Lors du traitement des 450 000 entrées en provenance du CBNA et d'Info Flora sur ArcGIS, le modèle tournait encore relativement rapidement. Cependant, plus les données étaient nombreuses plus les problèmes lors des sélections étaient fréquents. Les champs inutiles, et pratiquement jamais remplis, causaient aussi de nombreux *bugs* qui faisaient planter le modèle et afficher le message d'erreur 999 999 qui apparaît très souvent lorsqu'ArcGIS « doit gérer » de nombreuses données. Cependant, il arrive aussi qu'ArcGIS soit plus efficace que FME. Par exemple, le *transformer StatisticsCalculator* prend pas loin d'une heure pour effectuer son opération alors que son équivalent, le *Summary Statistics* d'ESRI, ne met qu'une soixantaine de secondes ! Joindre deux tables est également plus intuitif et également plus rapide au sein d'un *ModelBuilder* que sur FME. Probablement que la meilleure des solutions ne réside pas dans le choix d'un seul de ces deux logiciels mais en la combinaison des deux. Cela est désormais possible dès la version 10.3 d'ArcGIS.

Limitation de l'utilisation de la carte des milieux naturels pour la création de localités

La facilité de créer des prélocalités sur la base de la carte des milieux naturels est indéniable. Lier les notes à des groupements de milieux naturels permet de prendre en compte l'écologie des espèces et les prélocalités sont ainsi construites sur une base concrète. Cependant, il ne faudrait pas oublier que la carte est remise à jour chaque année et que certaines des notes ont plus de 10 ans. Il y a donc un biais à vouloir à tout prix inclure l'ensemble des notes au sein de groupements qui n'existent peut-être plus. De plus, la carte des milieux naturels, aussi précise soit-elle, est encore perfectible. Par exemple, les enrochements sur la rive droite du Rhône dans la région de Verbois n'apparaissent pas. L'autre « grand » problème est d'ordre topologique (**Figure 45**). Par exemple, pour le groupement N°1

(Végétation aquatique) la présence de ponts conduit à la création d'un ensemble de prélocalités complétement artificielles.

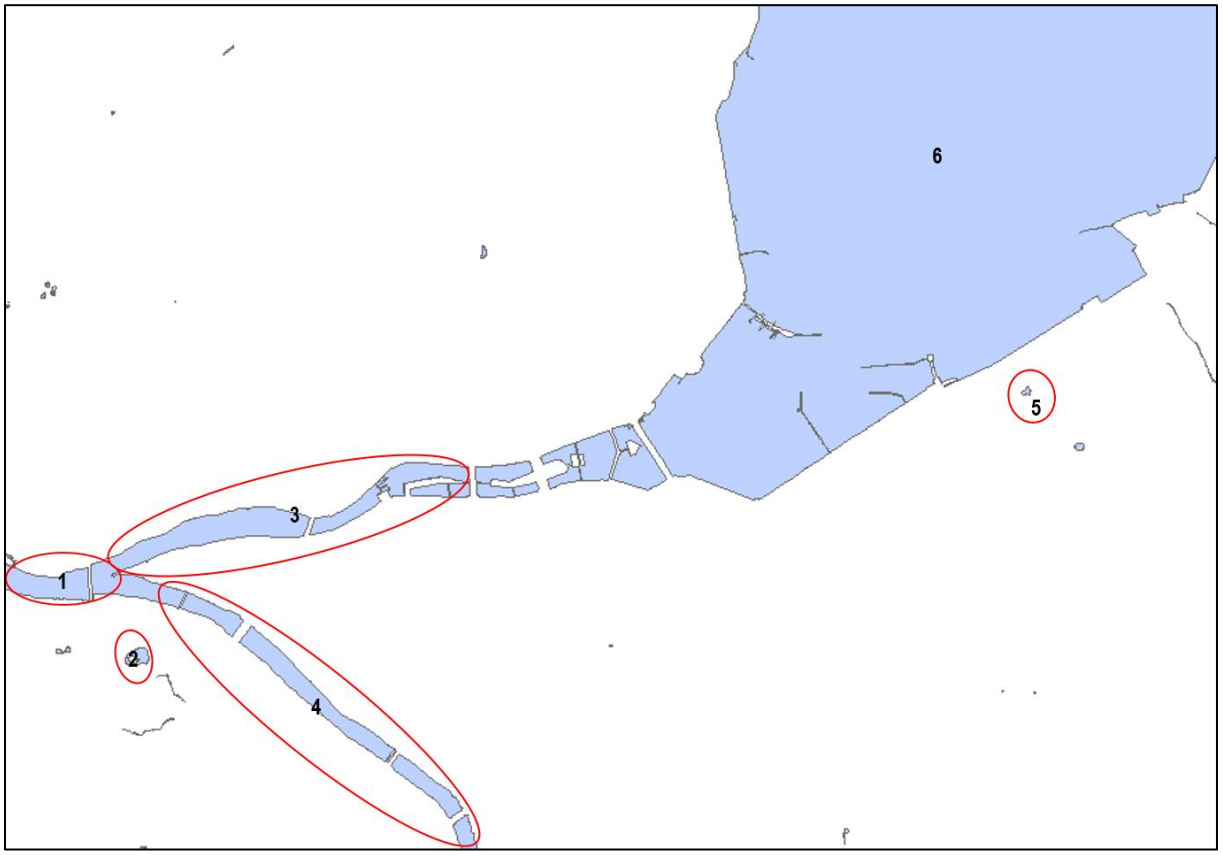


Figure 45 : Groupement N°1 (Végétation aquatique). Les ponts créent de nombreuses prélocalités artificielles. Les ellipses indiquent des localités créées à la fois sur la base de connaissances de terrain mais aussi sur le rendu visuel de la cartographie.

Dans l'exemple ci-dessus, la localité N°6 comprendrait l'embouchure du Rhône jusqu'au barrage du Seujet. Elle ne s'arrêterait donc pas au pont du Mont-Blanc. En revanche, la création d'une nouvelle localité est justifiée en aval du pont de la Jonction parce que la turbidité et les conditions physicochimiques (pH, carbonates, conductivité etc…) diffèrent des localités [4] et [3] respectivement l'Arve et le Rhône en aval du barrage du Seujet. Cet exemple montre une nouvelle fois qu'il est impossible de créer des localités de façon entièrement automatique sur la base des groupements de milieux naturels.

Limitation de l'utilisation des mailles pour le calcul de fréquence des taxons

Il arrive parfois que des notes tombent au carrefour de plusieurs mailles. Elles sont alors comptées à double. Il est relativement facile de retrouver ces notes et de supprimer les doublons. Mais comment choisir la maille qui devrait contenir la note ? Décaler l'ensemble du maillage est alors une solution. Un essai de décalage de 253 m a été effectué sur le maillage de 500 m. Le décalage de 253 m a été choisi afin d'éviter d'avoir des notes qui se retrouveraient à l'intersection de plusieurs mailles parce que leur précision aurait été arrondie. Pour certaines espèces, cela a induit un changement dans leur fréquence de répartition (AOO) qui pourrait être suffisant pour voir leur statut de menace évoluer. Il serait souhaitable de faire de nouveaux essais de décalages pour voir comment cela modifierait la fréquence de répartition des espèces et les conséquences que cela aurait sur leur statut de menace.

8. Conclusion

Bien que de nouveaux éléments aient été apportés sur la création des localités, il reste encore beaucoup de travail à effectuer. A terme, il serait souhaitable de transférer l'essentiel des modèles sur FME. Les buts du stage ont cependant été atteints. La création d'une méthodologie pour l'attribution d'un statut de menace et des modèles fonctionnels ont été créés. La génération des statuts de menaces a aussi été automatisée au maximum. Mais la création des localités et par extension leur comptabilisation ne pourront certainement jamais être totalement automatiques.

Quant au stage en lui-même, j'en suis très satisfait. Les nombreux problèmes rencontrés sur ArcGIS m'ont réellement aidé à progresser dans l'environnement ESRI. J'ai réellement appris plus que durant le premier semestre de ce certificat pour ce qui est de la programmation. En effet, l'utilisation de FME et du *ModelBuilder* ne totalisait que l'équivalent d'une matinée de cours. Il m'a donc fallut plonger rapidement dans le vif du sujet. Cela était d'autant plus crucial que le laps de temps accordé pour ce stage était très court. Peut-être aurait-il été souhaitable de commencer directement par FME afin de s'éviter ces petites frustrations qui ponctuèrent la création des différents modèles sur ArcGIS ? Dans tous les cas, ce stage aura permis de mettre le doigt sur les petits dysfonctionnements inhérents aux *ModelBuilder*. Je regrette cependant de ne pas avoir eu dès le début l'ensemble des données. Pouvoir faire quelques graphiques sur celles-ci auraient vraiment été un plus.

Dans tous les cas, le fait de pouvoir effectuer ce stage dans le cadre merveilleux des CJB a vraiment été une très bonne expérience. Celle-ci a été d'autant plus marquante que le stage s'inscrivait dans un projet concret qui aura des retombées sur la protection de la flore de notre région.

9. Bibliographie

Bornand C., Sager L. 2017. L'érosion rampante dans la diversité végétale. HOTSPOT. Vol 36. pp 6-7.

Bornand C., Gygax A., Juillerat P., Jutzi M., Möhl A., Rometsch S., Sager L., Santiago H., Eggenberg S. 2016. Liste rouge Plantes vasculaires. Espèces menacées en Suisse. Office fédéral de l'environnement, Berne et Info Flora, Genève. L'environnement pratique n°1621 :178 p.

Ceballos G., Ehrlich P.R., Dirzo R. 2017. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. Proceedings of the National Academy of Sciences, Early edition (DOI: 10.1073/pnas.1704949114).

Chevalier M. 2017. Liste Rouge des trachéophytes du canton de Genève, aide à la planification d'une campagne de terrain et évaluation des taxons. Travail de mémoire du Certificat complémentaire en géomatique. Université de Genève. 78 p.

Fischer M. et al. 2015. Etat de la biodiversité en Suisse en 2014. Forum Biodiversité Suisse et al. Berne. 92 p.

Gossner M.M et al. 2016. Land-use intensification causes multitrophic homogenization of grassland communities. Nature. Vol 540. pp 266-269.

Lacroix A. 2016. Régionalisation des critères UICN pour le canton de Genève dans le cadre de la réactualisation de la Liste Rouge de la flore vasculaire du canton de Genève. Conservatoire et Jardin Botanique. (Documentation interne). Genève. 33 p.

Lambelet-Haueter C., Schneider C., Mayor R. 2006. Inventaire des plantes vasculaires du canton de Genève avec Liste Rouge. Conservatoire et Jardin Botanique de la Ville de Genève. Genève.135 p.

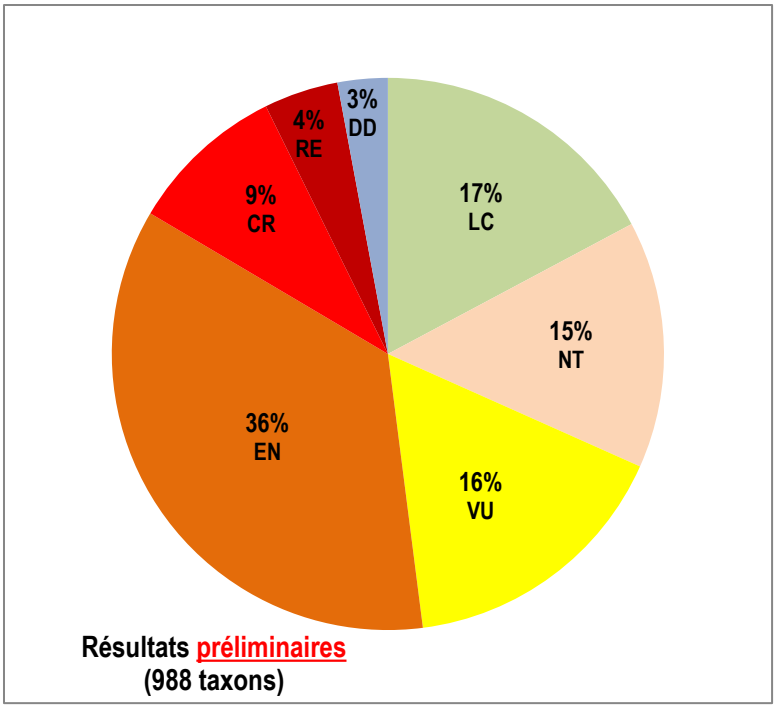
Theurillat JP., Schneider C., Latour C., Jeanmonod D. 2011. Atlas de la flore du canton de Genève. Catalogue analytique et distribution de la flore spontanée. Conservatoire et Jardin Botanique de la Ville de Genève. Genève. 720 p.

UICN France. 2011. Guide pratique pour la réalisation de Liste rouges régionales des espèces menacées-Méthodologie de l'UICN & démarche d'élaboration. Paris, France.

7. Annexes

Nom du groupement	CO_GRPMT	Correspondance avec les codes des milieux naturels genevois							
Végétation aquatique	1	101	102	107	203	218	10101	10102	10103
		10104	10105	10106	10107	10108	10109	10110	10111
		10112	10113						
Roselières et magnocaricaies	2	206	203	203	208	518	519	606	
Végétations amphibies	3	215	201	412					
Prairies humides et forêts humides ouvertes	4	210	211	212	213	214	411	615	623
		517							
Forêts inondables	5	604	606	605	607	501	614	518	519
Végétation pionnière sur substrat minéral remanié	6	301	302	518	520	904			
Milieux rudéraux et cultivés	7	201	201	208	208	208	208	208	215
		301	302	404	413	515	516	517	520
		619	701	708	801	802	803	806	807
		808	809	811	902	904	905	906	909
		913	999	999	1000	1002			
Zones minérales végétalisées	8	1005							
Forêts et ourlets mésophiles	9	501	502	502	516	517	517	601	602
		609	610	611	612	613	614	615	616
		619	620	1004					
Prairies et ourlets maigres / secs	10	213	404	406	407	410	413	416	502
		502	623	804	806	1002			
Prairies, gazons et ourlets eutrophes / mésophiles	11	407	804	603	1000	404	1006	1004	999
		808	413	403	807	904	515	510	913
		611	909	919	811	809	1003	502	411
		412	416	406	801	803	409	408	210
		212	410	517	201	201	902	701	1002
		805	806	906					

Annexe 1 : Table des correspondances entre les milieux naturels genevois et les groupements de milieux naturels



Annexe 2 : Statuts de menace provisoires de 988 taxons