

MARTIN SIROIS

**DÉVELOPPEMENT D'UN SYSTÈME
D'INFORMATION POUR LA PALÉOLIMNOLOGIE :
la base de données des diatomées circumpolaires
(Circumpolar Diatom Database - CDD).**

Mémoire présenté
à la Faculté des études supérieures de l'Université Laval
dans le cadre du programme de maîtrise en Sciences géographiques
pour l'obtention du grade de Maître en Sciences géographiques (M. Sc. Géogr.)

DÉPARTEMENT DE GÉOGRAPHIE
FACULTÉ DE FORESTERIE, DE GÉOGRAPHIE ET DE GÉOMATIQUE
UNIVERSITÉ LAVAL
QUÉBEC

2011

Résumé

Pour pallier le manque de connaissances sur la biogéographie des diatomées nordiques, le recueil d'un maximum d'informations est encouragé. Rassembler les données paléolimnologiques régionales permet d'effectuer des analyses écologiques et environnementales dans un contexte géographique plus large. La Circumpolar Diatom Database (CDD), un projet initié en 1997 par le Laboratoire de Paléoécologie Aquatique (LPA), tente de répondre à ces besoins. Toutefois, cette première version de la CDD comporte certaines lacunes, dont l'insertion de données de type « Z » (niveau de carottage) et l'implantation dans un système d'information géographique (SIG).

L'objectif principal de cette étude est de développer une nouvelle structure de base de données relationnelle particulièrement adaptée aux données paléolimnologiques favorisant l'archivage, la consultation et la représentation cartographique des données à l'échelle circumpolaire. En second lieu, ce projet vise à tester le potentiel d'implantation de la CDD dans un SIG et d'entrevoir sa diffusion au sein de la communauté scientifique.

Les étapes méthodologiques de développement d'un système d'information ont mené à la création d'un modèle de base de données relationnel composé par langage SQL. Une liste de références taxonomiques a été réalisée par l'unification de plusieurs listes existantes. Une interface utilisateur a été développée sous forme de formulaires MS Access facilitant la consultation des résultats. L'utilisation de la CDD a été testée dans ArcGIS par une connexion OLE DB et par conversion en « Personal Geodatabase ».

La collecte de données a permis de tripler le volume d'informations de la nouvelle CDD. Elle renferme maintenant quelque 572 sites d'échantillonnage, 39 350 données d'occurrence et 15 000 données limnologiques. Sa couverture s'étale sur huit régions circumpolaires arctiques réparties sur trois continents : Amérique du Nord, Europe et Asie. Jusqu'à maintenant, la couverture spatiale de la CDD permet difficilement d'effectuer des analyses sur une vaste échelle. Il est donc primordial d'encourager la communauté paléolimnologique à contribuer davantage au projet. Ce projet a généré un système d'information exploitable, et les phases d'implantation et de distribution pourront être entreprises en fonction de l'avenir souhaité pour la CDD et des ressources disponibles.

Remerciements

Je tiens tout d'abord à remercier mon directeur Reinhard Pienitz de m'avoir supervisé depuis la fin du baccalauréat jusqu'à ce mémoire de maîtrise. Je suis grandement reconnaissant de l'aide qu'il m'a apportée dans diverses facettes de mon cheminement, grâce à ses compétences, ses encouragements, sa confiance et sa compréhension.

Les précieux conseils et l'expertise de plusieurs ressources m'ont beaucoup été utiles pour les étapes de conception et de réalisation technique. Je remercie spécialement Étienne Girard de la Commission Géologique du Canada, Luc Cournoyer du Centre d'études nordiques, Éric Janssens-Coron du département de géomatique et Marc-André Bouchard du Ministère des Ressources Naturelles du Québec.

Également, je voudrais souligner la générosité et la confiance de tous les collaborateurs à la CDD, dont récemment Christian Bigler, Roland Hall, Peter Rosén et Dermot Antoniades qui ont contribué à améliorer la couverture de la CDD dans des régions où elle était peu représentée. Le soutien technique et moral d'une des instigatrices du projet, Marie-Andrée Fallu, m'a aidé à résoudre beaucoup de problèmes de vérification de données.

Merci à mon comité de supervision, Martin Lavoie et Marie-Hélène Vandersmissen de m'avoir fourni des commentaires constructifs et des encouragements pour la finalisation du projet.

Merci à toute l'équipe du LPA, de merveilleux collègues qui ont su m'accepter malgré l'excentricité de mon projet. Ma famille, mes proches qui m'ont compris et supporté à travers mes épreuves et, particulièrement ma compagne Geneviève Fradette pour m'avoir soutenu tout au long de mon parcours d'étudiant gradué.

Ce projet de recherche a été rendu possible grâce au soutien financier du programme ArcticNet par le biais de subventions accordées à Reinhard Pienitz, et le support du département de géographie et du Centre d'études nordiques de l'Université Laval.

Table des matières

Résumé	i
Remerciements	iii
Table des matières	v
Liste des tableaux.....	vii
Liste des figures.....	ix
Introduction.....	1
1. Chapitre 1 : L'étude des diatomées nordiques en paléolimnologie.....	5
1.1. Les données paléolimnologiques	5
1.2. Les diatomées	6
1.3. La région circumpolaire arctique	6
1.4. Les exemples connus de BD paléolimnologiques de diatomées	8
1.5. Historique du projet CDD	11
1.6. Problématiques et hypothèse	15
1.7. Objectifs de la recherche.....	15
2. Chapitre 2 : Les bases de données, systèmes d'information et systèmes d'information géographique	17
2.1. Définitions	17
2.2. Les modèles ou structures de base de données	19
2.3. Les avantages du modèle relationnel	20
2.4. Structure d'une base de données relationnelle.....	21
3. Chapitre 3 : Méthodologie du développement du système d'information	23
3.1. Les phases de développement générales.....	23
3.1.1. L'évaluation d'opportunité.....	23
3.1.2. L'analyse préliminaire	24
3.1.3. Architecture de système.....	24
3.1.4. Analyse fonctionnelle	25
3.1.5. Réalisation technique	26
3.1.6. Implantation	26
4. Chapitre 4 : Méthodologie spécifique à la CDD.....	27
4.1. Les modèles conceptuels du système d'information	27
4.1.1. Modèle conceptuel; structure préliminaire	28
4.1.2. Modèle conceptuel normalisé	29
4.1.3. Modèle final de base de données	31
4.2. Consultation par niveau de carottage.....	33
4.3. Les librairies de la CDD	35
4.3.1. Librairies de références générales.....	35
4.3.2. Librairies spécifiques à la CDD	37
4.3.3. Librairies taxonomiques	38
4.3.3.1. Sources des listes taxonomiques et des données de la CDD	39
4.3.3.2. Méthode de création de la librairie taxonomique	42
4.4. Importation des données	44
4.4.1. Uniformisation des unités	44
4.4.2. Disposition des tableurs	45
5. Chapitre 5 : Résultats	49

5.1.	Fichiers physiques de la CDD	49
5.2.	Sommaire du contenu de la CDD.....	50
5.3.	Description et fonctionnement des formulaires MS Access	51
5.4.	Tests d'utilisation dans un logiciel SIG	63
6.	Chapitre 6 : Analyse et interprétation des résultats	71
6.1.	Fonctionnement de la structure du SI.....	71
6.2.	L'interface utilisateur	72
6.3.	Bilan des différents tests d'utilisation	73
	Conclusion et recommandations	77
	Bibliographie.....	81
	Annexe A Liste détaillée des tables du fichier CDD.mdb	87
	Annexe B Liste détaillée des tables du fichier CDD_Taxa.Lib.mdb	127
	Annexe C Liste détaillée des tables du fichier CDD_Home.mdb.....	137
	Annexe D Code de programmation macro (MS Excel)	167
	Annexe E Codes de programmation VBA (MS Access)	171

Liste des tableaux

Tableau 1 : Les avantages des bases de données relationnelles.	21
Tableau 2 : Sommaire du contenu de la CDD (novembre 2010).....	51
Tableau 3 : Avantages et inconvénients de la connexion OLE DB.....	75
Tableau 4 : Avantages et inconvénients de la conversion en PersonalGDB.	75

Liste des figures

Figure 1 : Carte de la région circumpolaire-nord (NRCAN-ATLAS, 2009).....	7
Figure 2 : Localisation des sites de données paléolimnologiques incluses dans la Canadian Diatom Database (A) et la Diatom Paleolimnology Data Cooperative (B).....	10
Figure 3 : CDD FileMaker Pro, menu principal	12
Figure 4 : CDD FileMaker Pro, formulaire de résultats principal	13
Figure 5 : CDD FileMaker Pro, formulaire des régions	14
Figure 6 : CDD FileMaker Pro, formulaire de taxon.....	14
Figure 7 : Vue conceptuelle des phases de développement d'un système d'information	25
Figure 8 : Modèle conceptuel de la CDD; structure préliminaire.....	29
Figure 9 : Modèle conceptuel normalisé de la CDD.	30
Figure 10 : Modèle final de la CDD.	32
Figure 11 : Schématisation des variations d'épaisseur et d'intervalles des sous-échantillonnages en paléolimnologie.....	34
Figure 12 : Méthode de référencement des niveaux d'échantillonnage pour la CDD.....	35
Figure 13 : Exemple d'une déconcaténation de nom de taxon.	42
Figure 14 : Schématisation des interactions entre les listes taxonomiques de « CDD.mdb » et « CDD_Taxa_Lib ».....	44
Figure 15 : Transposition des données par niveau à l'aide d'une macro MS Excel.....	47
Figure 16 : Représentation des fichiers physiques de la CDD.	50
Figure 17 : Formulaire de menu principal « CDD Home page ».	52
Figure 18 : Extrait de la barre de menu latérale de CDD_Home.mdb.....	52
Figure 19 : Formulaire de recherche de résultats par lac « Results by lake ».	54
Figure 20 : Formulaire « Results by lake », onglet « Diatom Graph » (graphiques d'abondances relatives).....	55
Figure 21 : Formulaire « Results by lake », onglet « Water Chem » (résultats d'analyses d'eau).	56
Figure 22 : Formulaire « Results by lake », onglet « LOI » (perte-au-feu).....	57
Figure 23 : Formulaire « Results by Diatom ».	59
Figure 24 : Formulaire « Publication details ».	61
Figure 25 : Formulaire « People » onglet détails.....	62
Figure 26 : Formulaire « People », onglet publications.....	63
Figure 27 : Détail de l'utilisation de la connexion OLE DB dans ArcGIS.	66
Figure 28 : Détail de l'utilisation de la CDD convertie en Personal Geodatabase dans ArcGIS.	68
Figure 29 : Profondeurs des lacs de la CDD.....	69
Figure 30 : Mesures de la transparence de l'eau par disque Secchi.	70

Introduction

Dans la majorité des domaines de recherche en géographie, la question de l'archivage et la gestion des données accumulées devient tôt ou tard un problème qui doit être solutionné. En paléolimnologie, une science visant à reconstituer les conditions lacustres passées par des indicateurs biologiques et physico-chimiques, on n'échappe pas à cette réalité. Pourtant, il n'existe encore que peu d'exemples de base de données répondant aux besoins de cette science, même que les exemples connus sont peu ou pas adaptés à la paléolimnologie, surtout en ce qui concerne les régions nordiques.

Par ailleurs, les jeux de données et résultats paléolimnologiques issus de projets régionaux sont rarement synthétisés et interprétés dans un contexte géographique plus large. L'immensité du territoire et les coûts associés aux déplacements rendent difficiles les études dans les régions arctiques et expliquent en partie pourquoi ces travaux régionaux sont rarement combinés pour couvrir un plus vaste territoire.

Au Centre d'études nordiques (CEN) de l'Université Laval, le Laboratoire de Paléoécologie Aquatique (LPA) a pressenti ces besoins à combler en matière de regroupement de données et leur archivage. Le LPA est un laboratoire qui se spécialise dans les reconstitutions environnementales à l'aide d'études paléolimnologiques. L'un des principaux bioindicateurs utilisés pour ces études est la diatomée. C'est en 1997, après avoir cumulé de nombreux résultats sur plus de sept ans de travaux de recherche, que le LPA a entamé son premier projet de base de données de diatomées circumpolaires, la Circumpolar Diatom Database (CDD). À ce moment, le projet visait à rassembler les données pour fins d'archivage, de consultation et d'analyse. En plus des données sur les occurrences diatomifères, cette première version de la CDD renfermait des informations sur les caractéristiques limnologiques (température, pH, matière organique totale, chlorophylle *a*, carbone organique dissous, conductivité, concentration d'éléments chimiques et isotopiques, etc.) actuelles et passées. Cette première version de la CDD n'a toutefois pas été développée selon le modèle de base de données entités-association, ce qui ne permet pas l'insertion de données de type « Z » (différents niveaux d'analyse d'un même carottage). De plus, cette première structure de base de données ne permettait pas son implantation dans un système d'information géographique (SIG).

Or, avec les récents avancements dans le domaine des systèmes d'information géographique, de nombreuses possibilités d'analyses sont désormais plus accessibles aux chercheurs. La puissance des outils SIG en matière d'analyses spatiales et statistiques permet d'effectuer des requêtes complexes plus aisément qu'auparavant. Que ce soit grâce à des représentations cartographiques précises, des analyses statistiques ou des analyses spatio-temporelles, les outils SIG peuvent fournir d'importants atouts pour l'interprétation des données paléoécologiques. La paléolimnologie, étant un champ de recherche qui fait partie de la biogéographie¹, tirerait profit à utiliser davantage le potentiel des outils SIG pour l'approfondissement des connaissances sur les diatomées, entre autres, pour l'étude des aires de répartition en relation avec les facteurs environnementaux.

L'aire de répartition de nombreuses espèces diatomifères de régions arctiques demeure encore inconnue. Plusieurs études ont démontré l'existence d'espèces de diatomées arctiques endémiques, mais l'aire de répartition de nombreuses autres espèces reste à déterminer. C'est pourquoi l'incorporation d'un maximum d'informations écologiques et environnementales relatives aux taxons circumpolaires s'impose (Bouchard *et al.*, 2004; Vyverman *et al.*, 2007).

Les informations contenues dans la CDD peuvent alimenter l'interprétation des patrons de répartition des diatomées par rapport à des facteurs environnementaux actuels et passés. L'importance de mettre à jour l'ancienne structure de cette base de données pour la rendre plus adaptée aux outils SIG est donc préconisée. Les possibilités que permettrait le développement de ce système d'information sont nombreuses : archivage, consultation, accessibilité, sécurité, diffusion réseau, standardisation, validation, traitements statistiques, partage, extraction, représentations cartographiques, analyse spatio-temporelle, métadonnées, auteurs, sources et bien d'autres.

Le principal objectif de ce travail consiste à développer une nouvelle structure de base de données particulièrement adaptée aux données paléolimnologiques, favorisant la consultation et la représentation cartographique à l'échelle circumpolaire. Par ailleurs, ce

¹ La biogéographie est une science qui étudie la répartition des espèces vivantes à la surface du globe, ses causes et ses modifications (Parent, 1990).

projet vise à tester le potentiel d'implantation de la CDD dans un SIG et de diffusion au sein de la communauté scientifique.

Pour réaliser ce projet, une démarche méthodologique issue à la fois de la paléolimnologie et de la géomatique a dû être suivie. Le parcours entrepris a nécessité la maîtrise de champs de connaissances distincts, soit celui de l'étude des diatomées nordiques et celui des systèmes d'information et bases de données. Afin d'adapter le contenu de ce mémoire aux lecteurs ciblés, quelques notions et concepts se devaient d'être expliqués d'entrée de jeu. Conséquemment, ce travail se divise en deux principales parties, soit les définitions et concepts, puis la méthodologie et les résultats. D'ailleurs, outre un chapitre décrivant les différentes étapes de développement d'un système d'information, un chapitre additionnel de méthodologie spécifique au projet précède le chapitre de résultats.

L'analyse préliminaire pour le développement de ce système d'information a fondé le choix du système de gestion de base de données (SGBD) comme étant Microsoft Access. Le développement de l'architecture de système a été réalisé conjointement dans Enterprise Architect et MS Access, menant à la sélection du modèle conceptuel final et des dictionnaires associés (librairies). La méthode choisie pour l'incorporation des données à plusieurs niveaux (Z) est basée sur le concept des bases de données de forages géologiques. Une liste taxonomique de diatomées de référence a été réalisée par l'unification de plusieurs listes existantes, dont la North American Diatom Ecological Database (NADED) List of taxa et la liste d'Umeå University (Suède), ainsi que six autres listes spécifiques issues d'études paléolimnologiques arctiques. La phase de réalisation technique et l'implantation ont été réalisées dans MS Access. Différents formulaires de consultation ont été développés afin de faciliter la navigation dans les résultats. L'implantation de la CDD dans ArcGIS a été testée par une connexion OLE DB et par importation dans une « Personal Geodatabase ». Finalement, des recommandations et perspectives d'avenir sont proposées en conclusion.

1. Chapitre 1 : L'étude des diatomées nordiques en paléolimnologie

1.1. Les données paléolimnologiques

D'un point de vue spatial, les travaux de recherche paléolimnologiques peuvent être divisés en deux grandes catégories ; les travaux régionaux et les travaux étendus. Les travaux régionaux visent à dresser des reconstitutions environnementales pour un seul ou quelques lacs dans un même territoire. Ce type de travaux nécessite le prélèvement de carottes longues à plusieurs niveaux de sous-échantillonnages. Quant aux travaux étendus, ils visent plutôt à échantillonner un grand nombre de lacs dispersés sur des territoires différents, souvent dans le but d'établir des étalonnages ou des fonctions de transfert. Les travaux étendus requièrent des prélèvements de carottes courtes à niveau unique de trois centimètres (cm) d'épaisseur ou moins. Dans chacun des cas, un ou plusieurs échantillons d'eau seront prélevés pour tous les lacs étudiés. De plus, un portrait des caractéristiques physiques et chimiques du site étudié sera effectué lors de la visite.

Quel que soit le type de recherche paléolimnologique, plusieurs analyses peuvent être effectuées sur les échantillons récoltés. Pour les analyses d'eau, fournissant des informations sur les conditions actuelles du lac, des analyses biologiques (phytoplancton (Chla), microcrustacés, etc.), physiques (acidité, conductivité, salinité, transparence, température, etc.) et géochimiques (nutriments, gaz, métaux lourds, etc.) peuvent être réalisées. Pour les analyses de sédiments, fournissant des informations sur les conditions passées (récentes et/ou anciennes) du lac, des analyses biologiques (diatomées, microcrustacés, pollens, pigments, etc.), géochimiques (nutriments, gaz, perte au feu (LOI), isotopes stables, polluants, datations, ions majeurs/mineurs, métaux lourds, etc.) et physiques (stratigraphie, granulométrie, varves, etc.) peuvent également être réalisées.

1.2. Les diatomées

Les diatomées (Bacillariophyta) sont des algues unicellulaires microscopiques (de 2 µm à 1 mm) autotrophes possédant un exosquelette siliceux nommé frustule. Elles constituent l'un des groupes algaux les plus abondants dans les écosystèmes aquatiques et représentent un maillon très important de la chaîne alimentaire. Werner (1977) suggère que les diatomées contribuent à la production primaire mondiale dans une proportion de 20 à 25 %. Les diatomées sont reconnues pour leur grande sensibilité face à des fluctuations dans les conditions limnologiques et environnementales. Par exemple, des changements au sein des communautés biologiques sont rapidement observés lors de modifications des conditions physico-chimiques ou de température. Leur frustule est conservé dans les sédiments permettant d'effectuer des analyses paléolimnologiques pour lesquelles les diatomées sont souvent utilisées comme bio-indicateurs afin de dresser des reconstitutions paléoenvironnementales et climatiques.

Plusieurs facteurs influencent la répartition des diatomées. De ceux-ci, trois grandes catégories sont particulièrement étudiées en biogéographie des microorganismes : les conditions environnementales, l'historique environnemental postglaciaire et les interactions biologiques (Bouchard *et al.*, 2004).

1.3. La région circumpolaire arctique

Le territoire d'étude de cette recherche s'étant dans la région circumpolaire-nord (Figure 1). La région circumpolaire-nord correspond à la région englobant l'Arctique et le Subarctique. Dépendant des définitions utilisées et des critères de division physiographique, la limite sud de cette région peut varier entre le 50^{ième} et le 65^{ième} parallèle nord. Dans le cadre de ce projet, la frontière subarctique de la région circumpolaire n'est pas fermée afin de favoriser la cueillette de données des régions mitoyennes qui sont fonction intégrante de l'écosystème circumpolaire. La région circumpolaire arctique compte plusieurs biomes de régions froides, dont les déserts polaires et les calottes glaciaires, la toundra et la forêt boréale. Les écosystèmes arctiques sont connus pour être à la fois vulnérables et résilients. Ces dernières décennies, d'importants changements climatiques sont survenus dans

l'Arctique et ceux-ci bouleversent autant les milieux naturels que les communautés (ACIA, 2005). L'étude des écosystèmes arctiques est privilégiée dans plusieurs domaines de recherche, notamment en paléolimnologie. En effet, les paysages des régions arctiques et subarctiques sont reconnus pour être particulièrement riches en écosystèmes d'eau douce (Vincent & Pienitz, 1996). De plus, les écosystèmes lacustres circumpolaires constituent d'excellents témoins des changements climatiques, puisqu'ils sont très sensibles aux variations environnementales (Pienitz *et al.*, 2004).



Figure 1 : Carte de la région circumpolaire-nord (NRCAN-ATLAS, 2009)

1.4. Les exemples connus de BD paléolimnologiques de diatomées

Certaines initiatives de création de bases (BD) de données diatomifères ont eu lieu au cours des dernières décennies, notamment en Europe et aux États-Unis. Parmi celles-ci, on en trouve qui sont essentiellement constituées de banques d'images, ayant pour but principal d'offrir un outil d'identification taxonomique. C'est le cas de la Klaus-Dieter Kemp's Diatom Database, de la collection de Friedrich Hustedt Diatom centre et de la Diatom Collection (California Academy of Sciences). Ces bases de données sont de très utiles et impressionnantes collections, mais leur structure est conçue sous forme de catalogue, ce qui ne permet pas aisément l'insertion des données dans un SIG pour l'analyse spatiale. Outre les informations qu'elles contiennent au sujet des publications et de la nomenclature se rattachant à chaque taxon, ces bases de données contiennent pourtant de nombreuses informations complémentaires, telles que les caractéristiques du site d'échantillonnage incluant souvent les coordonnées géographiques. Mais, la rigueur de collecte est avant tout concentrée sur les informations taxonomiques, ce qui ne permet pas d'arriver aux mêmes fins que celles visées par le projet de la CDD.

Il existe toutefois des exemples de bases de données de diatomées qui présentent des modèles de structure inspirants pour une intégration éventuelle des données dans un SIG : la Diatoms of Western North America, la Diatom Paleolimnology Data Cooperative (DPDC) (Sullivan & Charles, 1994) et la European Diatom Database (EDDI) (Battarbee *et al.*, 2001). Ces grandes bases de données offrent, en plus du nécessaire en matière de taxonomie, des données spatiotemporelles sur les diatomées qui sont assemblées et harmonisées avec les conditions environnementales associées à leur localisation. Grâce à ce travail de regroupement d'informations biologiques et environnementales, il a été possible d'établir ou de préciser des fonctions de transfert², notamment dans le cas de la European Diatom Database (Criticula, Newcastle University, Catalogue). Il existe également des bases de données désirant couvrir l'ensemble du territoire nord-américain comme la Canadian Diatom Database de la Commission géologique du Canada et la DPDC de

² Une fonction de transfert est une représentation mathématique de la relation entre la présence d'une espèce de diatomée et certaines conditions limnologiques influentes actuelles. Grossièrement, elle s'effectue par la comparaison d'assemblages de diatomées fossiles avec ceux de diatomées modernes pour lesquelles les conditions environnementales prévalentes sont connues.

l'Academy of Natural Sciences (Sullivan & Charles, 1994), mais la répartition des sites qui y sont répertoriés dévoile une faible couverture pour la région circumpolaire (Figure 2). Il existe également des initiatives plus récentes visant à cataloguer les résultats de recherche des territoires polaires, tels que le Polar Data Catalogue (ArcticNet) et Antarctic Freshwater Diatoms (University of Colorado). Certains projets de regroupement de données paléoenvironnementales pour fins d'analyses ont servi d'inspiration pour le projet de la CDD, tels que Predicting Aquatic Ecosystem Quality using Artificial Neural Networks (PAEQANN) (Gosselain *et al.*, 2005), Paleoecological Investigation of Recent Lake Acidification (PIRLA) (Charles & Whitehead, 1986; Whitehead *et al.*, 1990) et Surface Water Acidification Project (SWAP) (Munro *et al.*, 1990). La North American Pollen Database (NAPD) (Grimm, 2000), constitue également un exemple inspirant pour ce projet. En effet, la NAPD est issue d'un regroupement de plusieurs bases de données régionales. Elle a récemment été fusionnée avec son analogue européen, la European Pollen Database (EPD), pour former la Global Pollen Database (GPD) (Grimm, 2000). La GPD représente un exemple de réussite de projet d'unification de données biogéographiques et paléoenvironnementales.

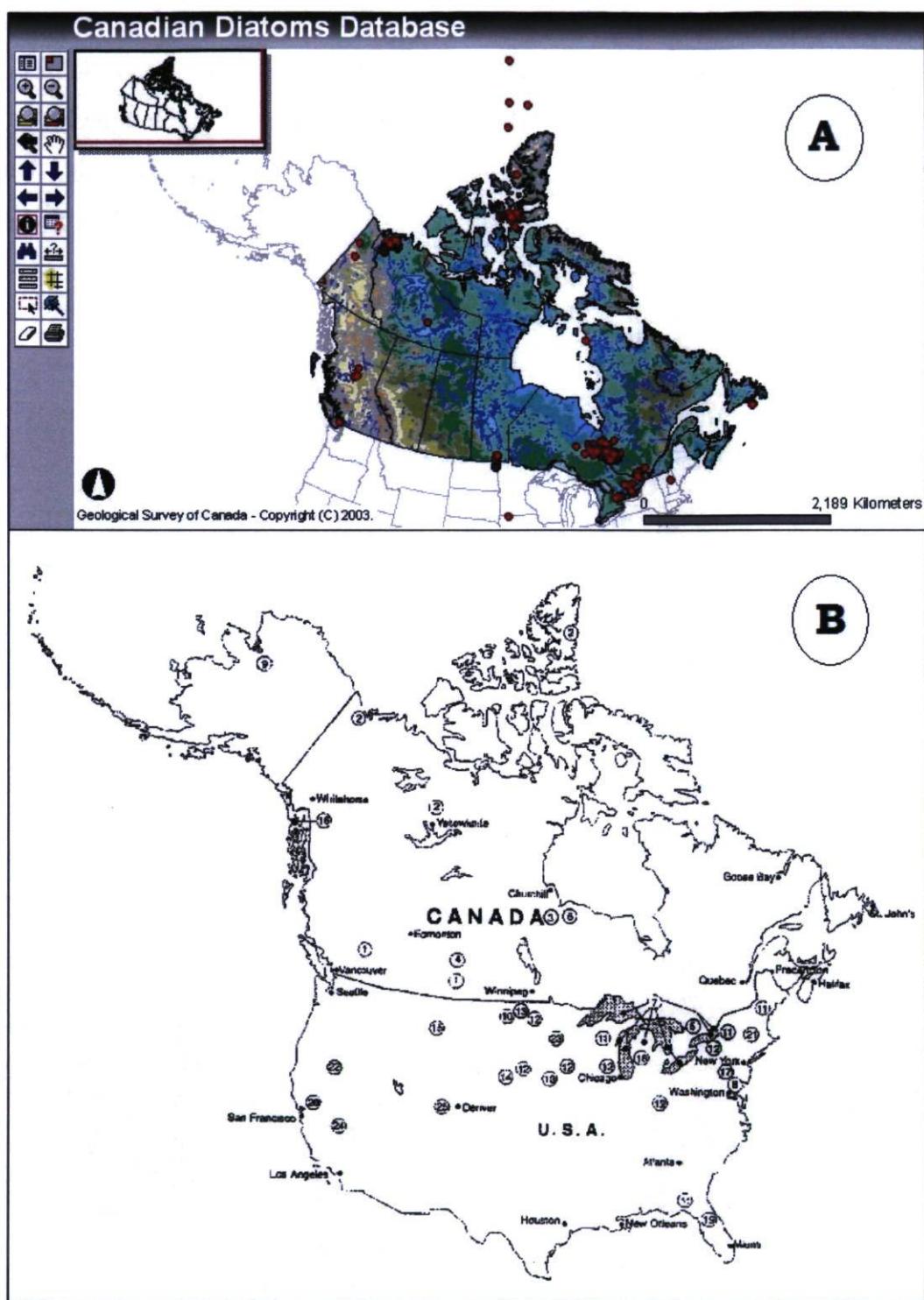


Figure 2 : Localisation des sites de données paléolimnologiques incluses dans la Canadian Diatom Database (A) et la Diatom Paleolimnology Data Cooperative (B).

1.5. Historique du projet CDD

La base de données des diatomées circumpolaires, ou Circumpolar Diatom Database (CDD), est un projet développé et supervisé par le Laboratoire de Paléoécologie Aquatique (LPA) au Centre d'études nordiques (CEN) de l'Université Laval depuis 1997. Les besoins d'accessibilité et d'archivage des données ont incité le démarrage de ce projet. La base de données vise à rassembler les nombreux résultats de recherches issues de travaux paléolimnologiques, tels que les occurrences diatomifères et les caractéristiques limnologiques. À l'origine du projet, la CDD comptait déjà 261 sites d'échantillonnage (lacs et étangs), couvrant cinq régions circumpolaires arctiques, soit le Québec nordique (Nunavik), le Labrador, le Nunavut, le Yukon et la Sibérie. Des résultats d'analyses physicochimiques étaient associés à chacun de ces sites ainsi que des données d'abondance relative de diatomées. 560 taxons ont été cumulés dans la banque taxonomique de cette première version de la CDD. La base de données est consultable à partir du logiciel FileMaker Pro (v.3.0 ou supérieur) et elle comporte cinq formulaires de consultation. Le premier formulaire est le menu d'accueil (Figure 3), qui contient le titre et l'affiliation de la base de données, ainsi qu'un bouton d'action « Entrer » pour commencer la consultation. Le deuxième formulaire est celui des principaux résultats (Figure 4), celui-ci réunissant des informations et relevés à propos de la région d'étude, les publications associées, les analyses d'eau et l'abondance relative d'un taxon sélectionné. Un menu horizontal en son entête permet d'accéder aux autres formulaires ou tables de la CDD. Le troisième est un formulaire d'informations sur les régions d'étude (Figure 5), comportant parfois une localisation géographique sommaire et les publications associées à la région sélectionnée. Le quatrième est le formulaire de taxon (Figure 6) affichant les abondances relatives pour tous les lacs où le taxon a été observé. Finalement, quelques vues des données en mode liste sont offertes à l'utilisateur ainsi que quelques fonctions de tri et de recherche.

Bien que toutes ces fonctionnalités permettent de naviguer assez aisément à travers les données, cette première version de la CDD n'a pas été développée selon le modèle de base de données entités-association, ce qui ne permet pas l'insertion de données de type « Z », soit les niveaux de carottage de la méthode d'analyse paléolimnologique. De plus, cette vieille structure de base de données ne permettait pas son implantation dans un système

d'information géographique (SIG). Ces limitations sont dues au fait que la première version de la CDD avait été conçue selon une structure de données dite « en liste ». C'est-à-dire que les données étaient essentiellement listées dans une seule table, provoquant des redondances d'information dans plusieurs champs. Cette structure a l'avantage d'être simple de conception, mais elle comporte toutefois plusieurs désavantages. En plus de ceux précédemment cités, on compte la mauvaise gestion de doublons et de l'intégrité des résultats, la difficulté de gestion et de mise à jour des données abondantes, ainsi que les limitations lors de la manipulation des résultats (n'utilisant pas le langage SQL pour les requêtes).

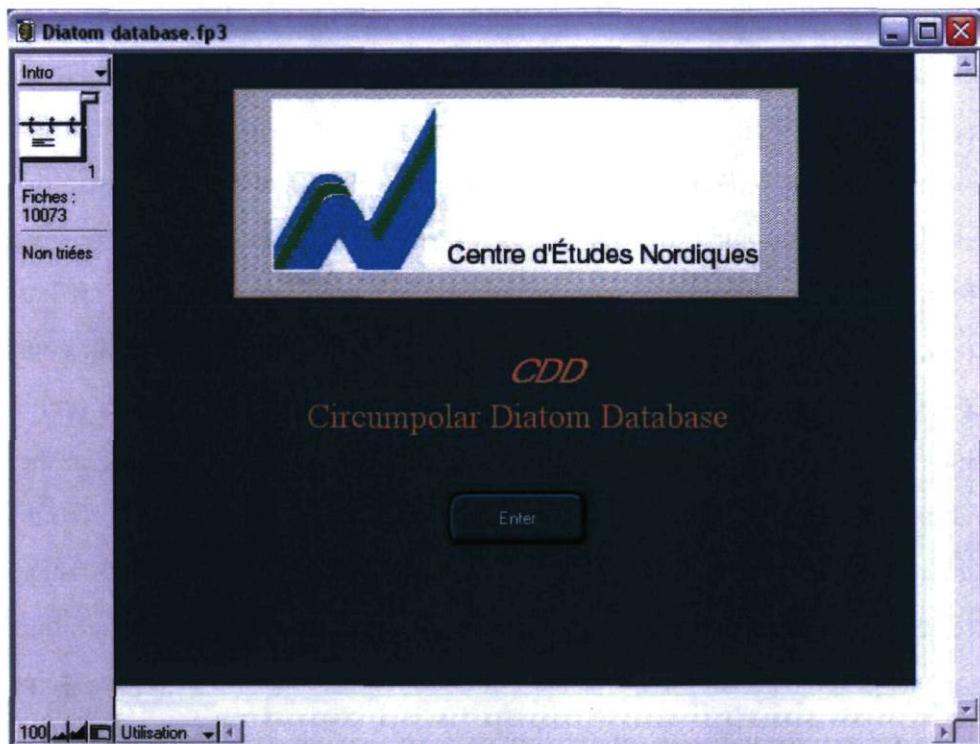


Figure 3 : CDD FileMaker Pro, menu principal.

CDD
Circumpolar Diatom Database

		Region	Code	Taxa	
Fiches :	10073	Yukon	YUK	Cocconeis placenta var. euglypta	
		Lake ID	Senior author	Taxon code	Occurrence (%)
		YUK001	R. Pienitz &	COCPLEU	18.21
Geographical setting		Physical variables	Nutrients	Major ions (mg/l)	Metals (µg/l)
LAT (N) 60.39		DEPTH (m) 32	TPU (µg/l) 14.1	Ca 21.5	Fe 8.4
LONG (W) 134.57		TRANS (m) 9.1	TPF (µg/l) 11	Mg —	Mn 5
ALT (m) 671		CLR (Pt units) —	P04 (mg/l) —	Na 46.3	Al —
AREA (ha) 181.2		TEMP (°C) 17.5	SRP (µg/l) 2.2	K 10.1	ZOOPL 25.3
CATCH (ha)		TEMP ADJ (°C) —	TN (µg/l) 559	Cl 63.6	CHL a (µg/l) 0.3
DIST (km) —		O2 (µg/l) 13.7	TKH (µg/l) 309	S04 300.2	CHL s (µg/l) 0.1
FOREST 1		pH 8.7	HCO3 (µg/l) 0.3	SiO2 2.81	Co —
WOODLAND		ALK (µg/l) —	NO3 (µg/l) 208	Cr —	—
TUNDRA		COND (µS/cm) 700	NH3 (µg/l) 12	Cu —	—
COASTAL TUN		ORG (g) —	PN (µg/l) 42	Li —	—
ALPINE TUN		—	DOC (mg/l) 10.6	Mo —	—
—		—	DIC (µg/l) 88.2	Ni —	—
—		—	POC (µg/l) 321	Pb —	—
—		—	POC/PN 7.6	Sr —	—
—		—	TN / TP 39.6	V —	—
—		—	—	Zn —	—

Figure 4 : CDD FileMaker Pro, formulaire de résultats principal.

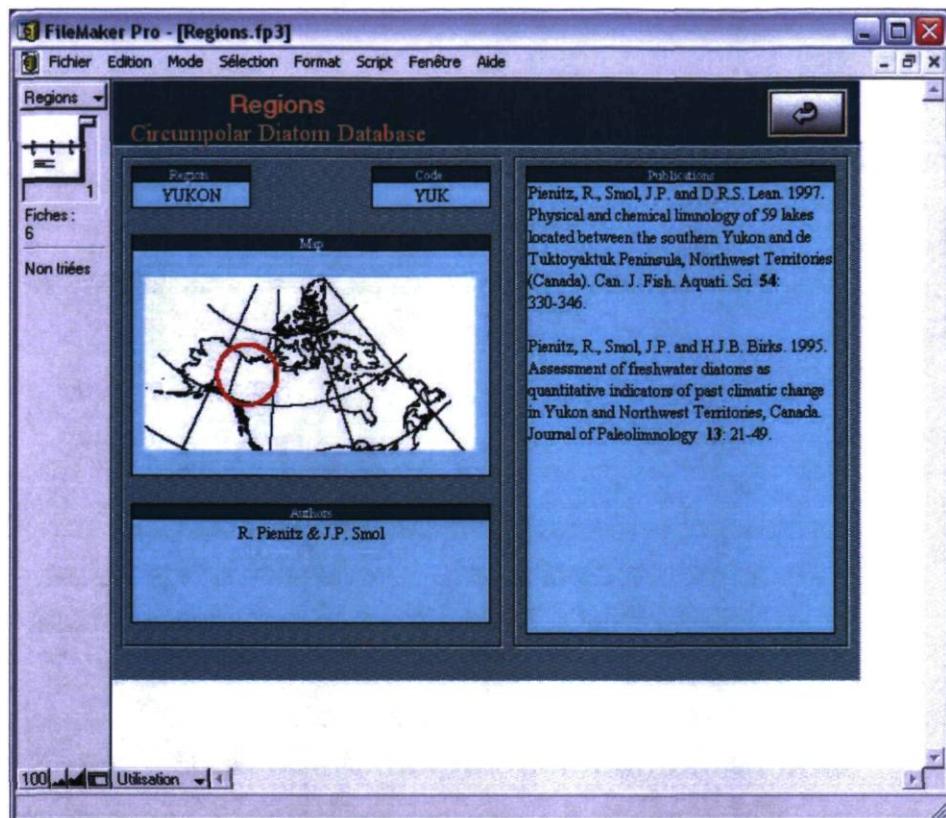


Figure 5 : CDD FileMaker Pro, formulaire des régions.

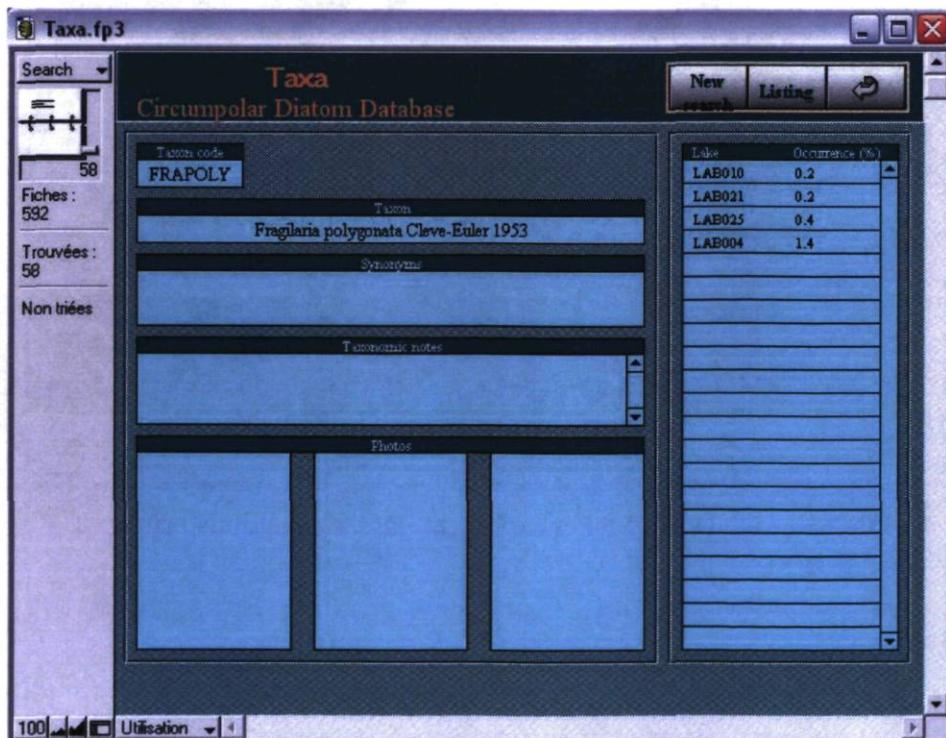


Figure 6 : CDD FileMaker Pro, formulaire de taxon.

1.6. Problématiques et hypothèse

En paléolimnologie, aucun outil adéquat n'est à la disposition des chercheurs pour archiver et consulter les résultats d'études de régions circumpolaires arctiques. Or, pour pallier le manque de connaissances sur la biogéographie des diatomées nordiques, le recueil d'un maximum d'informations est encouragé.

Nous supposons que le développement d'une base de données de type relationnelle adaptée aux données paléolimnologiques pourrait permettre d'instaurer un système d'information allouant l'archivage, la consultation et le traitement des résultats. Ces développements pourraient permettre l'implantation du système dans un logiciel de cartographie assisté par ordinateur SIG.

1.7. Objectifs de la recherche

L'objectif principal de cette étude est de développer une structure de base de données relationnelle particulièrement adaptée aux données paléolimnologiques favorisant l'archivage, la consultation et la représentation cartographique des données à l'échelle circumpolaire. Les principaux objectifs secondaires sont d'évaluer le potentiel de la CDD pour son implantation dans un SIG et d'entrevoir sa diffusion au sein de la communauté scientifique.

2. Chapitre 2 : Les bases de données, systèmes d'information et systèmes d'information géographique

Avant d'expliquer les méthodes qui ont permis d'obtenir les résultats de cette étude, quelques définitions et concepts de base sont ici exposés. Il s'agit d'apporter certaines précisions concernant les différentes sphères de travail touchées par ce projet. En effet, certaines notions de géomatique et des systèmes d'information tirent avantage à être éclaircies avant d'être élaborées plus en détail dans les prochains chapitres.

2.1. Définitions

Base de données

Une base de données est un ensemble structuré de données pouvant servir dans différentes applications (Bergeron, 1993). L'abréviation de base de données est BD et en anglais DB pour « database ». En d'autres mots, c'est une entité dans laquelle il est possible d'entreposer des données de façon structurée, généralement sous forme de tables, et organisées selon certains critères permettant l'exploitation de ces données tout en limitant leur redondance. Dans une BD, les données peuvent servir à des programmes informatiques, et ce, par différents utilisateurs simultanément. Lorsqu'il y a ainsi partage des données, on parle alors d'un réseau, notion qui est souvent associée à celle de base de données. Le Grand Dictionnaire Terminologique (OQLF, 2010) relate que les bases de données sont parfois confondues avec les banques de données. Ces dernières sont plutôt des ensembles d'éléments d'information propres à un domaine défini et rendus disponibles à plusieurs utilisateurs.

Système d'information

En général, ce que l'on appelle système d'information désigne l'ensemble de la structure développée pour collecter, classer, archiver, gérer, traiter et partager les données. Le système d'information (SI) est donc constitué d'un ensemble d'équipements, de procédures, de ressources humaines et de données qui y sont traitées, ayant pour objectif de divulguer de l'information à différents utilisateurs d'une ou plusieurs organisations (Bergeron, 1993).

Système d'information géographique

Un système d'information géographique (SIG) est un système d'information qui réunit des données descriptives et des données géométriques géoréférencées. Les données géométriques sont gérées comme des objets géographiques dans un SIG. Les objets géographiques et les données descriptives sont interconnectés dans un SIG, de sorte qu'il est possible d'interroger le système d'un côté comme de l'autre de ces groupes de données (Thériault, 1996a; Bergeron, 1993). Comme le SI, le SIG est un ensemble constitué de plusieurs ressources matérielles et humaines visant à partager des informations, auquel s'ajoutent les notions de présentation et d'analyse de données géoréférencées.

Système de gestion de base de données

Un système de gestion de base de données (SGBD), en anglais DBMS (Database management system), est un système logiciel qui permet de gérer une base de données de façon manuelle ou automatique. Il permet d'opérer les tâches de saisie, de stockage, de traitement, de mise à jour, de diffusion et de protection des données (Bergeron, 1993). Le SGBD gère la BD sur plusieurs niveaux, de la gestion des fichiers sources, au traitement interne des données jusqu'à l'interface utilisateur.

Langage de requête SQL

Le SQL (Structured Query Language) est un langage de requête structuré, standardisé par l'Organisation internationale de normalisation (ISO), qui permet de manipuler, définir ainsi que d'interroger des données. Le langage SQL est étroitement lié aux bases de données relationnelles, pour lesquelles il est depuis longtemps un standard. La majorité des SGBD supportent le langage SQL. Néanmoins, des versions étendues et propriétaires du langage SQL de base peuvent être utilisées par différentes compagnies de SGBD (Codd, 1990; Bisson, 1997).

Normalisation

La normalisation est une tâche propre au développement d'un modèle de base de données, donc la phase d'analyse de système. Elle consiste à définir une nouvelle relation entre deux objets d'un modèle de base de données dès qu'une relation 1:N ou N:N est possible. Elle

permet de s'assurer que les valeurs des différents attributs soient bien en dépendance fonctionnelle avec la clé primaire (Thériault, 1996b; Roche, 2009). La dénormalisation suit le plus souvent la normalisation. Celle-ci consiste à retirer les relations superflues entre des objets pour lesquels des relations 1:N ou N:N sont moins nécessaires, même si cela crée quelques répétitions de données ou des champs additionnels.

Librairies

Dans une base de données, une librairie, aussi appelée table de référence ou dictionnaire, est un catalogue normalisé qui inventorie et définit les éléments de données d'un système d'information (Bergeron, 1993).

2.2. Les modèles ou structures de base de données

Plusieurs types de structures peuvent être utilisés pour la conception d'une base de données. Leurs différences résident principalement dans la structure des données qu'elles contiennent, également appelée modèle de représentation des données. Parmi les structures de base de données les plus communes, on compte la structure hiérarchique, la structure réseau, la structure relationnelle, la structure multidimensionnelle et la structure orientée-objet.

Le modèle hiérarchique

Le modèle hiérarchique est le premier modèle de données qui a été développé. Sa structure a la forme d'un classement hiérarchique descendant. Des pointeurs informatiques sont utilisés pour relier entre eux les enregistrements contenus dans la base de données.

Le modèle réseau

Le modèle réseau s'apparente fortement à la structure ramifiée du modèle hiérarchique. Des pointeurs sont également utilisés entre les enregistrements, sans toutefois se restreindre aux liens descendants entre ceux-ci.

Le modèle relationnel

Le modèle relationnel, aussi appelé le modèle entité-association, est composé de plusieurs tables en deux dimensions dans lesquelles les données sont listées sous forme de lignes (enregistrements) pouvant posséder de nombreuses colonnes (champs). Les données sont consultables par le biais de la théorie mathématique des relations. Cette structure de base de données non hiérarchique permet d'établir des liens entre les données dans plusieurs sens possibles. Le modèle relationnel est celui qui est le plus utilisé dans le monde depuis la fin des années 1990 (Hainaut, 2009 ; Roche, 2009).

Le modèle multidimensionnel

Le modèle multidimensionnel est une variation du modèle relationnel auquel s'ajoute une troisième dimension. Chaque cellule incluse dans les tables tridimensionnelles, aussi appelées cubes ou hypercubes, contient des données agrégées associées aux entités de chacune de ses dimensions.

Le modèle orienté-objet

Le modèle orienté-objet organise les données sous forme de classes dans lesquelles se trouvent des champs en tant qu'instance de classe. Cette structure de base de données permet de gérer aisément différents types d'objets, tels que des graphiques, des images, du contenu audio, du texte et d'autres types de données. Ce modèle est souvent employé pour gérer des données géoréférencées contenant des objets de nature surfacique, vectorielle et matricielle, ainsi que pour les environnements multimédias et la conception Internet.

2.3. Les avantages du modèle relationnel

Les fondements théoriques des bases de données relationnelles ont été développés en 1970 par Edgar Frank Codd, un chercheur britannique en sciences informatiques travaillant pour la compagnie IBM. Ses travaux ont eu une influence majeure dans le domaine des sciences informatiques, de telle sorte que le modèle relationnel est le modèle de base de données le plus utilisé depuis les années 1990. De nos jours, il existe encore plusieurs avantages à l'utilisation d'un modèle relationnel pour la création d'un système d'information. Outre

l'utilisation des standards SQL comme langage d'interrogation et de manipulation de données ainsi que de ses compatibilités réseau, de nombreuses caractéristiques favorisent le choix du modèle relationnel pour le développement d'un système d'information tel que la CDD (Codd, 1990). Le Tableau 1 en énumère quelques-uns.

Tableau 1 : Les avantages des bases de données relationnelles.

<u>Avantages</u>	<u>Descriptions</u>
Adaptabilité	plus facile et moins prompt à l'erreur lors d'une modification à la structure d'une table et/ou donnée.
Atomicité des opérations	possibilité d'annuler des modifications sans perte de données.
Compatibilité d'ouverture	possibilité de lire la base de données à partir d'un autre programme utilisant le standard SQL (ex. logiciel de cartographie, interface Internet, logiciel statistique, etc.).
Sécurité d'investissement	les relations étant basées sur un fondement logique, il est improbable que la méthode soit remplacée d'ici peu de temps.
Validation des données	validations automatiques lors de l'insertion ou de la modification de données.
Domaines de valeurs	restriction du contenu possible dans un champ.
Indépendance logique des données	possibilité de diviser une table en deux sans en affecter les données.
Indépendance physique des données	possibilité de diviser une base de données en plusieurs fichiers sur un ou plusieurs systèmes.
Indépendance d'intégrité	possibilité de modifier les relations et dépendances des données sans en affecter les données et les fichiers.
Indépendance de distribution	possibilité d'utiliser un programme externe pour manipuler les données grâce au langage SQL.
Toujours en ligne	possibilité d'ajouter et/ou de modifier des données, des formulaires, rapports, etc. sans avoir à verrouiller la base de données.

2.4. Structure d'une base de données relationnelle

Une base de données relationnelle est constituée de classes d'objets (tables) interconnectées par des associations (relations). Une classe d'objets peut se décrire comme un fichier tabulaire composé d'enregistrements (objets) et de champs (attributs), s'organisant respectivement en lignes et en colonnes. La valeur associée à un attribut d'un

enregistrement constitue une donnée, aussi appelée occurrence. Les relations entre les tables d'une base de données relationnelle sont établies grâce aux identifiants (clés).

Plusieurs types de relations peuvent exister entre les objets d'une base de données, les plus fréquentes sont les relations un à un (1:1), un à plusieurs (1:N) et plusieurs à plusieurs (N:N). Les relations 1:1 sont établies lorsque deux objets possèdent une majorité d'attributs communs. Par exemple, un lac et ses coordonnées X et Y sont en relations 1:1, puisqu'un lac ne peut posséder plusieurs coordonnées et que les coordonnées ne peuvent référer à un autre lac. Lorsqu'une relation 1:1 est établie, les objets sont placés dans la même table de la base de données.

Quant aux relations 1:N, celles-ci sont implantées lorsqu'il peut exister plusieurs occurrences pour un même objet. Par exemple, un lac peut avoir été échantillonné plus d'une fois, donc plusieurs carottes de sédiments peuvent être issues d'un même lac. Dans ce cas-ci, l'objet « Carotte de sédiments » doit être placé dans une autre table. Pour qu'une association physique soit établie entre les deux classes d'objets, une clé primaire doit être attribuée à « lac » (un) et une clé étrangère doit être insérée dans la table de « Carotte de sédiments » (plusieurs). La clé étrangère correspond à un champ ajouté dans une table. Ce champ est une copie de l'identifiant unique de la clé primaire, à la différence que les répétitions sont autorisées pour permettre les relations 1:N.

Une relation N:N est établie lorsque deux objets peuvent être associés plusieurs fois l'un à l'autre. Par exemple, un collaborateur peut être membre de plusieurs affiliations et une affiliation peut posséder plusieurs membres. Dans de tels cas, une nouvelle table doit être créée, celle-ci comprendra les clés primaires de chacun de ces objets sous forme de clés étrangères interconnectées. La table ainsi créée se nomme table de concordance, ou encore table de relation (Thériault, 1996a-b; Hainaut, 2009 ; Roche, 2009).

3. Chapitre 3 : Méthodologie du développement du système d'information

Pour déterminer les étapes méthodologiques de ce projet de recherche, les phases de développement d'un système d'information publiées par le gouvernement du Québec (Ministère des Communications, 1993) ont été consultées. Comme c'est le cas pour la plupart des méthodes de développement de systèmes d'information, une série de principes de base ont été considérés pour la méthodologie de ce projet : la construction de modèles, du général au particulier, du conceptuel au physique, une solution d'ensemble, un système de qualité et un système en évolution (Thériault, 1996a; Roche, 2009). Ces principes constituent de multiples étapes de planification préalables à la conception de la BD et à l'instauration du SI. Dans le cadre du présent projet, le développement de la CDD s'est déroulé en six phases : l'évaluation d'opportunité, l'analyse préliminaire, l'architecture de système, l'analyse fonctionnelle, la réalisation technique et l'implantation. Au cours de ce chapitre, chacune de ces étapes de développement est décrite de façon générale. Quant au chapitre suivant, il décrira plus en détail les étapes spécifiques qui ont été réalisées dans le cadre de ce projet.

3.1. Les phases de développement générales

3.1.1. L'évaluation d'opportunité

Cette première phase de développement consiste à cerner les problèmes à résoudre, à explorer les solutions possibles pour les corriger, puis faire le rapport entre les besoins et les ressources matérielles et monétaires disponibles. Cela permet de déterminer les principaux objectifs à atteindre et de définir la meilleure solution technologique qu'il sera possible de réaliser en fonction des investissements disponibles. Dans le cas présent, les investissements correspondent à un projet de maîtrise, donc environ deux ans et demi de travail. Une fois l'évaluation d'opportunité réalisée, il est alors jugé opportun ou non d'aller de l'avant vers le stade de l'analyse préliminaire.

3.1.2. L'analyse préliminaire

Cette phase permet d'élaborer davantage les solutions technologiques nécessaires à la correction des problèmes soulevés lors de l'évaluation d'opportunité. Il s'agit donc ici de préciser l'étendue et les causes des problèmes observés, développer des scénarios de solutions et les mettre en perspective en les confrontant aux contraintes de temps et d'investissements. De par les éléments qu'elle fournit, l'analyse menée à cette phase du développement répond aux questions soulevées vis-à-vis de la validité, la faisabilité, des investissements nécessaires, des avantages et des impacts liés au développement du système d'information.

3.1.3. Architecture de système

D'un point de vue conceptuel, cette troisième phase de développement d'un système d'information, autant que les deux phases subséquentes, peuvent être synthétisées comme le passage du niveau conceptuel au niveau physique. En effet, telles qu'illustrées à la Figure 7, ces trois étapes se résument au « Quoi », qui réfère au niveau conceptuel, au « Comment » référant au niveau logique et, finalement, la « Réalisation » correspondant au niveau physique.

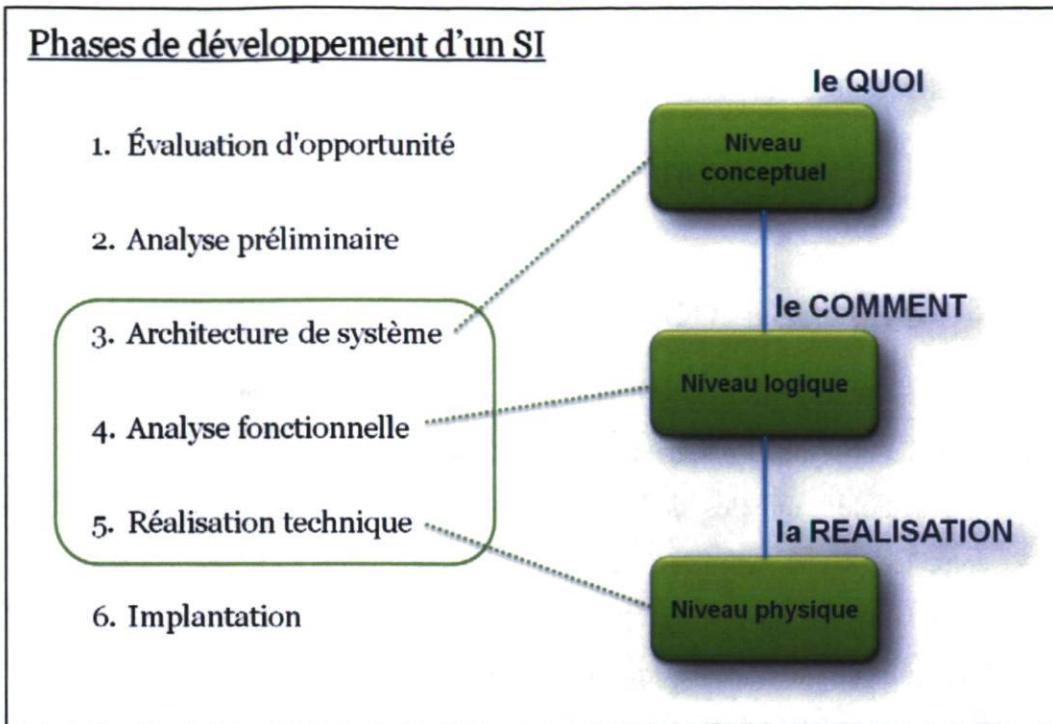


Figure 7 : Vue conceptuelle des phases de développement d'un système d'information.

La phase d'architecture de système a comme principal objectif de planifier et modéliser le système d'information. Préalablement au développement de la structure, il y a d'abord la sélection des grandes fonctionnalités que le système devra mettre à disposition des utilisateurs, ainsi que le choix du type d'interface utilisateur approprié. Une fois les fonctionnalités et l'interface sommairement déterminées, la création du modèle conceptuel de base de données peut débuter. Le modèle conceptuel permet de développer la structure de données qui sera utilisée par le système.

3.1.4. Analyse fonctionnelle

Ce stade de développement, référant au « Comment » dans la vue conceptuelle (Figure 7), consiste à parachever le modèle conceptuel et de concevoir le modèle logique de base de données. Il s'agit d'adapter le modèle aux particularités du projet, telles que l'organisation spécifique des données complexes et les sources de références, aussi appelées dictionnaires ou encore librairies. Les librairies permettent de référer à des informations récurrentes dans plusieurs tables. Elles ont pour fonction d'alléger le volume de la base de données tout en la

rendant plus uniforme et sécuritaire en évitant les redondances et en diminuant les risques d'erreurs. Dans certains cas, les librairies peuvent se référer à des listes approuvées ou universelles, ce qui permet également de rendre les données comparables avec d'autres systèmes d'information qui emploient les mêmes références.

3.1.5. Réalisation technique

Le stade de la réalisation technique consiste à matérialiser les modèles conceptuels et logiques en des composantes matérielles. C'est à cette étape que les enregistrements sont insérés dans les tables, que les définitions des champs sont inscrites, et que les liens physiques entre les différentes clés primaires et secondaires sont appliqués concrètement. Une série d'essais et de tests d'interrogation et de validation des données sera exécutée au cours de l'insertion des données et de la création du système. Les différents modules de l'interface utilisateur sont également conçus à cette étape.

Préalablement à l'insertion des données dans la structure finale, des méthodes d'uniformisation, de traitement et de restructuration des données sources doivent être établies et appliquées dans un tableur de données. Une fois la structure physique concrétisée et les données sources conformes aux standards d'importation, la base de données peut alors être remplie. Cette étape fastidieuse s'effectue par importation de fichiers textes, sous forme de tableurs, ainsi que par requêtes SQL spécifiques à chaque type d'enregistrement. Ces opérations sont décrites en détails au chapitre 4.

3.1.6. Implantation

La dernière phase de développement d'un système d'information est l'implantation, aussi appelée la mise en service. C'est ici que s'effectuent les derniers ajustements pour préparer la diffusion du système. La diffusion peut se faire de plusieurs façons, entre autres de manière physique (support matériel), par un réseau privé ou un réseau public. Des essais finaux doivent être appliqués au système avant diffusion. Des documents informatifs et la formation de ressources humaines peuvent être nécessaires en fonction de l'envergure du projet.

4. Chapitre 4 : Méthodologie spécifique à la CDD

Tel que décrit à la section 1.4, la première version de la CDD comptait plusieurs désavantages que ce projet tente d'éliminer. Les premières étapes de développement consistaient donc à effectuer une évaluation d'opportunités et une analyse préliminaire. Celles-ci ont permis d'établir la validité d'aller de l'avant et de développer une structure relationnelle pour la prochaine version de la CDD. Quant aux phases suivantes de conception du système, elles ont mené à la création de plusieurs outils de développement qu'il faut préciser ici. Parmi les plus importants, on compte : la volumétrie projetée, les modèles de système, la structure de données par niveau de carottage, la sélection et création des librairies, l'annexion d'une base de données taxonomique, le protocole de standardisation, les patrons de restructuration de données sources et les requêtes d'importation en langage SQL. Le présent chapitre explique en détails cette méthodologie spécifique qui a été utilisée pour répondre aux particularités du projet Circumpolar Diatom Database.

4.1. Les modèles conceptuels du système d'information

Aux cours des phases d'architecture système et d'analyse fonctionnelle, la création de modèles de structure relationnelle pour la base de données connaît une évolution constante jusqu'à l'atteinte du modèle conceptuel final qui sera employé pour ce projet. D'abord, il s'agit d'établir les grands ensembles d'objets qui constitueront la base de données et d'établir les principales relations entre ceux-ci. Par la suite est dressée une première énumération exhaustive des champs qui seront contenus dans chacun de ces ensembles. Progressivement, les schémas conceptuels seront normalisés à chaque essai, rendant le modèle toujours plus complet et avec une meilleure logique relationnelle entre les données. Tel que décrit au chapitre 2, la normalisation consiste à créer un nouvel ensemble (table) à chaque fois qu'une relation 1:N est possible, afin de s'assurer que les valeurs des différents attributs soient bien en dépendance fonctionnelle avec la clé primaire. En résumé, le processus de conception de modèles conceptuels évolue du général au particulier, c'est-à-dire que les premiers modèles contiennent de grands ensembles et peu de relations, et que

les modèles subséquents contiennent de nombreux ensembles plus précis avec plus de relations entre eux (Roche, 2009 ; Thériault, 1996b).

4.1.1. Modèle conceptuel; structure préliminaire

Le premier modèle conceptuel de la base de données a réuni de façon plutôt hiérarchique les ensembles autour d'un élément de départ, c'est-à-dire les lacs. Ceux-ci ont été jugés comme étant le pivot central des relations logiques entre les données pour ce projet. À ce stade de modélisation, les champs contenus dans l'ensemble « LAKE » (lac) sont nombreux et contiennent principalement des informations de localisation, mais également des résultats de relevés *in situ*. Tels qu'illustrés à la Figure 8, deux autres ensembles viennent se greffer à celui des lacs par le biais de relations un à plusieurs (1:N). Il s'agit des deux principaux groupes de résultats issus d'analyses paléolimnologiques : les résultats d'analyses d'eau, nommés ici « Water Chemistry » et les résultats d'analyses de carottes de sédiments nommés « CORE ». À l'intérieur de ces deux ensembles se regroupe une multitude de champs, ceux-ci étant plus souvent associés aux types d'analyse, mais également au positionnement spatial dans le cas de la table « CORE ». Plus loin dans la structure relationnelle, un autre groupe de tables se déploie à partir de la table « CORE », il s'agit des résultats par niveau nommés ici « Niveau » ainsi qu'une table de caractéristiques « Caractéristiques_Niveau ». Finalement, trois tables subséquentes se rattachent à « Niveau » et contiennent les informations relatives aux espèces de diatomées. La table « Especies_Niveau » renferme les abondances pour chaque niveau et les tables « Especies » et « Especies_1 » contiennent une banque d'informations taxonomiques, la deuxième table permettant de créer des liens pour les synonymes taxonomiques.

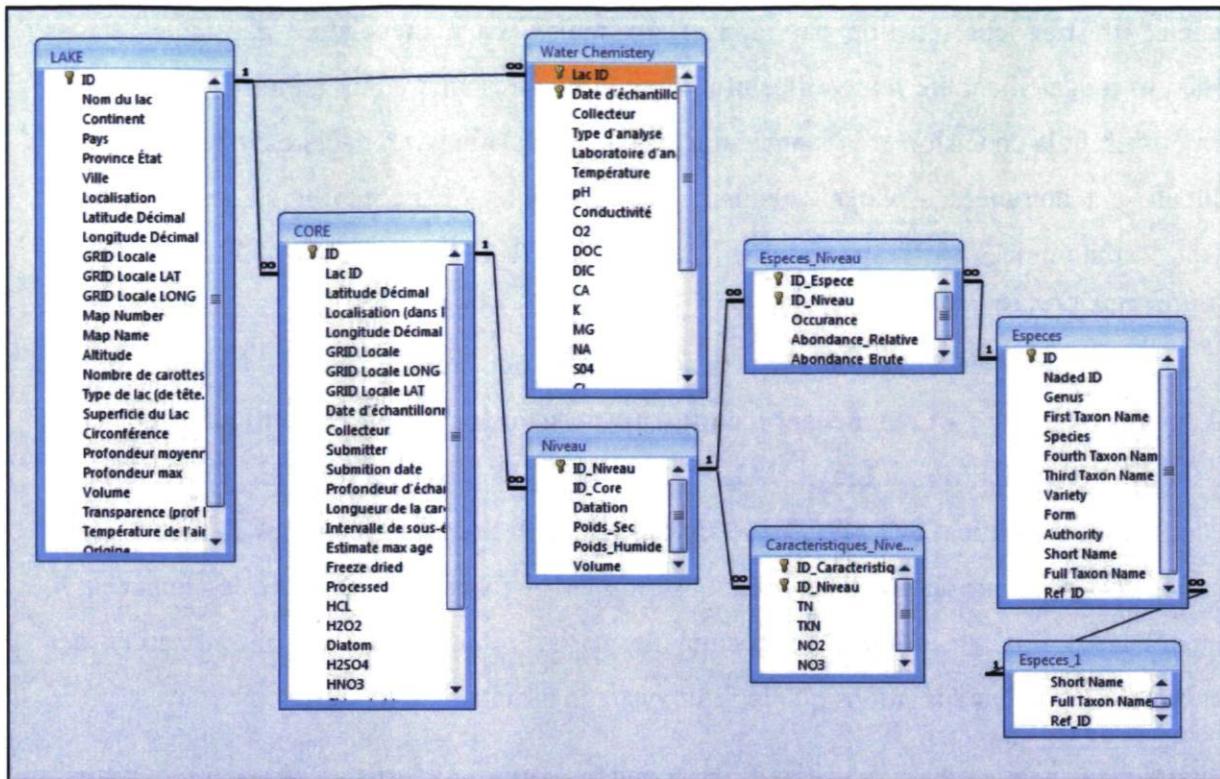


Figure 8 : Modèle conceptuel de la CDD; structure préliminaire.

4.1.2. Modèle conceptuel normalisé

Ce deuxième modèle (Figure 9) est basé sur la première ébauche de structure précédemment décrite et a été réalisé à l'aide du logiciel Enterprise Architect. Le modèle précédent contenait beaucoup d'éléments dans chaque ensemble puisqu'il visait à rassembler et à hiérarchiser les principaux objets et thématiques concernés par ce projet. Pour l'élaboration de ce deuxième modèle, un travail de normalisation a été effectué. L'étape de normalisation consiste à exploiter chaque relation un à plusieurs qu'il est possible de développer. Il en résulte la création de plusieurs nouvelles tables dont : « VISIT », « WATER_CHEM », « CORE_LEVEL », « ID_RESULT_CHEM », « REF », « ID_LEVEL_CHEM », « PEOPLE », « ID_LEVEL_COUNT », « AFFILIATION » et « LAB ». De plus, un premier ensemble de librairies a été incorporé au modèle, cet ensemble étant composé des tables : « Lib_COUNTRY », « Lib_UNIV », « Lib_ANALIT », « Lib_UNIT », « Lib_DIATOMS » et « Lib_SYNON ». Les librairies ont été représentées par des icônes de livres afin d'alléger la lecture du modèle et pour

mieux illustrer leur fonction par rapport aux tables. La nomenclature des tables et des champs a également été retravaillée afin de mieux correspondre aux thèmes référents, tout en priorisant la concision et la langue anglaise dans la chaîne de caractères. Ainsi, les tables auparavant nommées « Water Chemistry », « Niveau », « Espèces » et « Espèces_1 » ont été renommées respectivement « WATER_CHEM », « CORE_LEVEL », « Lib_DIATOMS » et « Lib_SYNON ». Autre fait notable, les tables « Espèces_Niveau » et « Caractéristiques_Niveau » du premier modèle ont été retirées. Le contenu de ces dernières tables se voit maintenant réparti dans plusieurs autres plus spécifiques, telles que « ID_LEVEL_CHEM », « Lib_ANALIT », « Lib_UNIT » et « ID_LEVEL_COUNT ». Le choix de ces modifications a été fondé sur la logique des relations entre les éléments, c'est-à-dire en s'assurant de créer des relations simples et fonctionnelles entre les données du modèle. Au final, le modèle conceptuel normalisé fournit une liste plus détaillée des ensembles et de leurs relations que le futur système d'information pourra contenir.

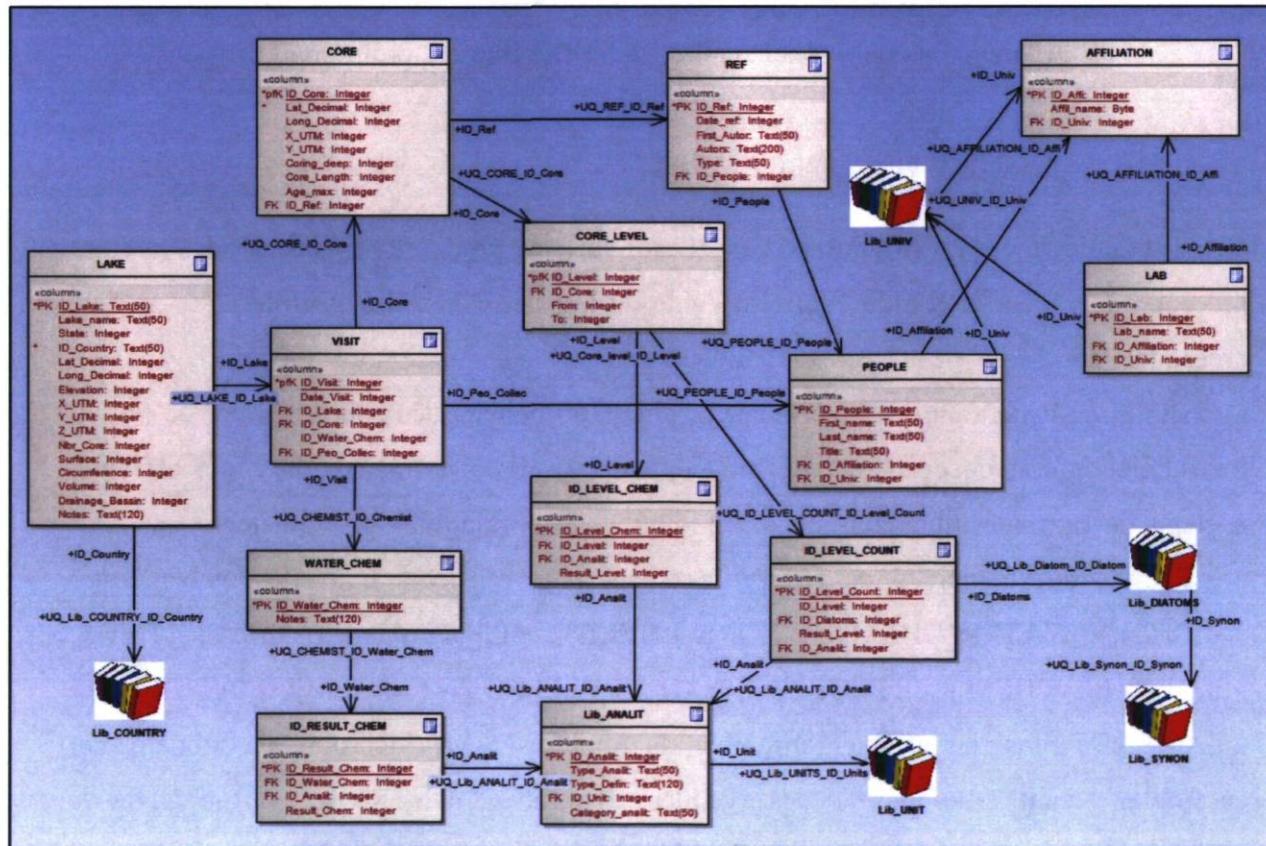


Figure 9 : Modèle conceptuel normalisé de la CDD.

4.1.3. Modèle final de base de données

Le modèle précédemment décrit tentait d'exploiter au maximum les relations un à plusieurs qu'il était possible de trouver. Au cours de la création du modèle final (Figure 10), un travail de dénormalisation a, entre autres, été effectué. Certaines tables ont ainsi été fusionnées et d'autres ont été créés dans le but de peaufiner la structure du modèle. Une vision plus technique a dû être adoptée lors de cette étape, dans le but d'assurer une bonne intégration du modèle dans le système logiciel. Conséquemment, des tables de concordance ont été créées lorsqu'un objet d'une table était questionné plusieurs fois. Ce type de tables contient deux clés primaires qui ont des relations récursives entre elles. Les tables de concordance permettent d'accéder aux informations de plusieurs tables de façon simple et sécuritaire (Codd, 1990). Dans le présent modèle, les tables de concordance qui ont été ajoutées sont « PEOPLE_AFFIL » et « RESULT_PUB ». En comparant le modèle final (Figure 10) avec le modèle normalisé (Figure 9), on constate que les principaux ensembles sont maintenant disposés en trois sections mieux définies : les résultats d'analyse en haut à l'horizontale, les métadonnées en bas à l'horizontale, et les librairies à l'extrême droite disposées verticalement. De plus, certaines tables du modèle précédent ont été abolies ou fusionnées, telles que : « ID_LEVEL_COUNT », « ID_LEVEL_CHEM » et « Lib_SYNON ». Certaines ont été renommées, telles que « ID_RESULT_CHEM » et « REF » maintenant « PUB » et « RESULT_WATER », respectivement. La nomenclature des tables et des champs a été parachevée en suivant la même méthode que pour le précédent modèle. De plus, de nouvelles tables ont été ajoutées : « SAMPLING », « PEOPLE_AFFIL », « RESULT_PUB » et « Lib_CLASS_RESULT ». Somme toute, le rendu final de ce modèle correspond à la structure physique réelle de la base de données. Les relations entre les données devraient permettre la consultation de l'entièreté du contenu une fois incorporé.

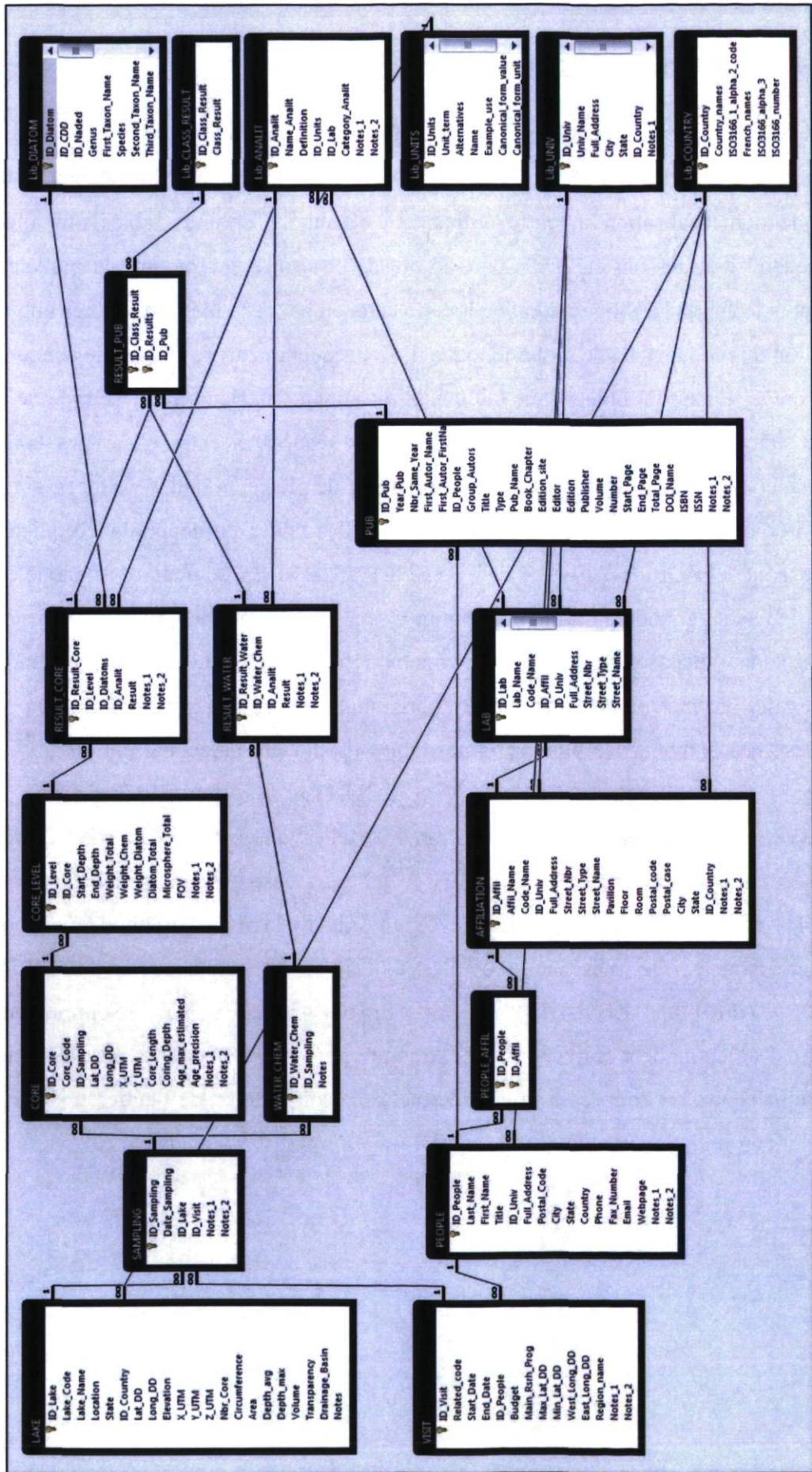


Figure 10 : Modèle final de la CDD.

4.2. Consultation par niveau de carottage

Un des besoins à combler dans le cadre de ce projet est la consultation des données par niveau de carottage. En paléolimnologie, les méthodes d'échantillonnage et de sous-échantillonnage peuvent varier d'un projet à un autre en fonction des objectifs de la recherche. Il n'est pas rare que pour une même carotte de sédiments, l'intervalle d'échantillonnage soit irrégulier (Figure 11). L'épaisseur des sous-échantillons peut également varier à l'intérieur d'une même carotte. Or, il est primordial que ces informations soient rattachées à chaque donnée qui provient d'un niveau de carottage, et ce, afin de pouvoir retracer à tout moment l'endroit dans la carotte où le sous-échantillon a été prélevé ainsi que son contexte d'échantillonnage. Il a donc été établi que trois critères doivent être pris en considération lors de l'importation de données de niveau de carottage : la profondeur, l'épaisseur et la distance par rapport à un autre sous-échantillon.

Au point de vue de la structure des données, ces informations multiples doivent être incorporées de façon uniforme et comparable. Les méthodes employées en géologie pour les bases de données de trous de forages ont été la principale source de référence pour l'élaboration de cette partie de la CDD (Commission Géologique du Canada). Dans ce type de base de données, les informations réfèrent à un niveau d'une carotte de roche dont la profondeur de départ est connue, l'épaisseur est variable, mais l'intervalle d'échantillonnage est continu (aucun saut dans le pas d'échantillonnage). Dans un tel cas, chaque niveau peut porter un identifiant de profondeur arbitraire, puis l'épaisseur dans un autre champ de la même table. Cependant, cet exemple ne peut s'appliquer dans le cas de sous-échantillonnage non-continu, comme c'est le cas en paléolimnologie. Il a donc été nécessaire de développer une technique légèrement différente, dont la démarche est décrite plus bas.

Initialement, il a été décidé que la profondeur de départ de 0 cm serait attribuée à la surface de la carotte. Ce principe stratigraphique est depuis longtemps appliqué dans les représentations graphiques de résultats paléoécologiques (Smol *et al.*, 2001; Pienitz *et al.*, 2004) et son efficacité logique s'applique également dans le présent contexte de structuration de données. Toujours selon ce principe, la profondeur dans la carotte est un

chiffre positif référant à la distance depuis la surface, celle-ci le plus souvent exprimée en centimètres avec deux décimales. Sur la base de ces principes, il a été possible d'élaborer une méthode d'archivage ou d'importation de données associées aux niveaux de carottage constituée de trois champs (Figure 12). D'abord, un premier champ est attribué pour l'identifiant unique de chaque enregistrement, ce qui est d'usage lors de la création d'une nouvelle table dans toute base de données relationnelle. Par la suite, la profondeur de départ du sous-échantillon est importée dans le champ « Prof_Départ ». Finalement, la profondeur de fin du sous-échantillon est insérée dans le champ « Prof_Fin ». Grâce à cette méthode d'importation, il est possible de connaître avec précision la profondeur, l'épaisseur et l'intervalle de chaque sous-échantillon à partir de ces deux seules informations de profondeur.

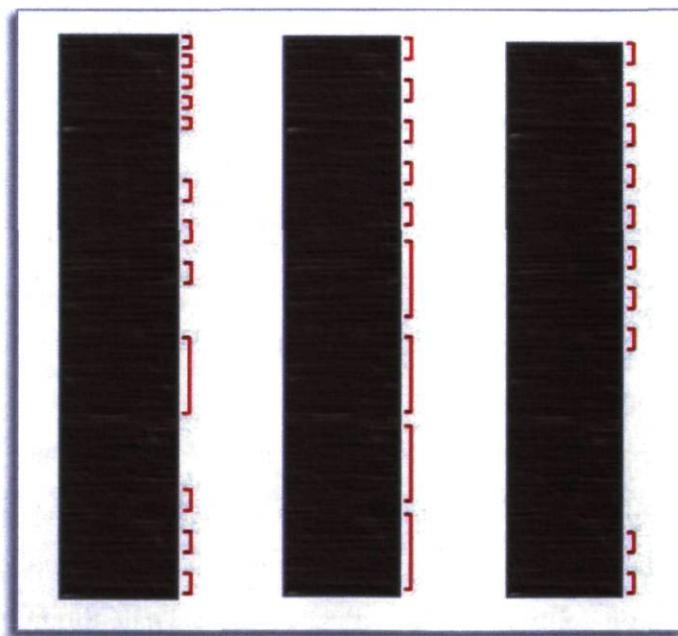


Figure 11 : Schématisation des variations d'épaisseur et d'intervalles des sous-échantillonnages en paléolimnologie.

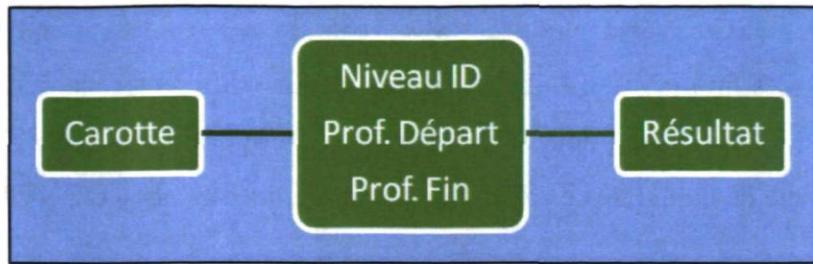


Figure 12 : Méthode de référencement des niveaux d'échantillonnage pour la CDD.

4.3. Les librairies de la CDD

Pour assurer l'uniformité des données, alléger le volume et diminuer le risque d'erreur ou de doublons au sein du système d'information, plusieurs listes de références ont dû être créées. Certaines ont été puisées à partir de sources existantes et d'autres ont été développées spécialement pour la CDD. Cette section de ce mémoire décrit les librairies qui ont été développées et leur méthode de création. Les huit librairies de la CDD peuvent être regroupées en trois catégories. La catégorie de références générales rassemble celles qui sont essentiellement issues de sources existantes. Une deuxième catégorie regroupe les librairies qui ont été développées en fonction des besoins spécifiques à la CDD. Finalement, la librairie taxonomique et ses sous-ensembles composent la troisième catégorie. De façon complémentaire à ces descriptions, un bilan détaillé pour chacune de ces librairies, ainsi que pour toutes les tables de la CDD, est disponible aux annexes A, B et C.

4.3.1. Librairies de références générales

Pays

Les résultats de recherches accumulés dans la CDD proviennent de plusieurs pays. De plus, les centres de recherche et les affiliations des collaborateurs de la CDD réfèrent parfois à différents pays. Afin d'éviter d'éventuels doublons ainsi que des fautes de frappe dans ces noms de pays, la pertinence d'une table de référence est justifiée. Nous nous sommes donc tournés vers l'emploi d'une liste de noms et de codes de pays basée sur les normes de l'Organisation internationale de normalisation (ISO). Une liste de 246 pays a donc été insérée dans la CDD. Celle-ci contient les noms anglais et français de chaque pays ainsi

que leurs codes ISO 3166-1-alpha-2, ISO 3166-alpha-3 et leur nombre ISO 3166 correspondant. Ces informations ont été tirées depuis différentes sources en ligne et rassemblées en une seule liste, dont ISO³, Wiki⁴ et CIA World Fact Book⁵. Dans la CDD, cette librairie issue du cumul de ces références a été nommée « Lib_COUNTRY ».

Universités

Au même titre que les pays, les universités et les centres de recherche sont cités à plusieurs endroits dans la base de données. Une librairie de noms d'universités mondialement reconnues a donc été insérée à la CDD. Cette liste contient 3040 noms d'universités et provient du Massachusetts Institute of Technology⁶. Les universités contenues dans cette liste ne représentent pas l'entièreté des institutions existantes dans le monde, mais elle s'avère suffisante pour les besoins de ce projet. À des fins de mise à jour éventuelle de cette librairie, d'autres champs ont été ajoutés à la table « Lib_UNIV ». Il s'agit de champs précisant l'adresse officielle des universités, le pays et la province où elles se situent. Ces champs permettront d'obtenir des informations plus précises sur les universités auxquelles les données de la CDD sont associées et ils pourront être mis à jour au fur et à mesure qu'une nouvelle université sera citée dans la CDD.

Unités

L'unité de mesure de chaque résultat d'analyse est une information de première importance en paléolimnologie, comme dans toutes les sciences quantitatives d'ailleurs. Relativement aux méthodes employées et aux normes régionales en vigueur, il arrive que les unités d'un même type d'analyse soient différentes d'un laboratoire à un autre. Les symboles unitaires peuvent également varier dans différents travaux, risquant d'occasionner des erreurs si importées telles quelles dans une base de données. La création d'une librairie d'unités composée d'une liste de noms et symboles d'unités de mesure s'avère donc essentielle dans le cadre de ce projet. En paléolimnologie, plusieurs unités de mesure relèvent d'analyse de concentration d'un élément en solution. La source initiale qui a servi à la création de cette librairie est issue de Health Level Seven (HL7, 2010). Cette référence répertorie une foule

³ http://www.iso.org/iso/fr/country_codes/iso_3166_code_lists.htm

⁴ <http://en.wikipedia.org/>

⁵ <https://www.cia.gov>

⁶ <http://stuff.mit.edu/people/cdemello/univ.html>

de noms et de symboles unitaires communément utilisés par « The Unified Code for Units of Measure » (UCUM) pour les concentrations en solution. Cette référence constituait une bonne base pour la création de la librairie de la CDD. Elle a ensuite été complétée par une douzaine d'unités de masse, de taille et de comptage qui sont fréquemment utilisées en paléolimnologie et qui étaient manquantes dans la liste de référence. La librairie résultante a été nommée « Lib_UNITS » et compte actuellement 234 enregistrements.

4.3.2. Librairies spécifiques à la CDD

Méthodes d'analyse

La librairie « Lib_ANALIT » a été créée entièrement pour la CDD. Elle est composée de diverses méthodes d'analyse qui ont servi à l'obtention de tous les résultats de la base de données, toutes catégories confondues. Elle contient une liste d'abréviations⁷ communément utilisées dans la littérature scientifique, leur description et la catégorie d'analyse associée. Deux relations sont rattachées à cette table, soit une vers la librairie « Lib_UNITS » et l'autre vers la librairie de laboratoires d'analyse. Dans le futur, « Lib_ANALIT » pourra éventuellement être enrichie de nouvelles méthodes, mais pour le moment elle en compte 92.

Laboratoires

La librairie « LAB » regroupe les laboratoires associés autant aux travaux de terrain, aux méthodes d'analyse des résultats qu'aux collaborateurs de la CDD. On y répertorie les noms, abréviations, acronymes et adresses de laboratoires. De plus, les librairies affiliations et universités sont jointes à « LAB » par des relations 1:N. Jusqu'à maintenant, six laboratoires sont listés dans cette librairie. De par son faible nombre d'enregistrements et sa grande probabilité à connaître des mises à jour, l'abréviation « Lib_ », normalement utilisée pour référer aux librairies, n'a pas été ajoutée à « LAB ». Cette attention préviendra le gestionnaire du système d'information que cette table de référence est davantage sujette aux modifications que d'autres librairies.

⁷ Par exemple, les abréviations pH, DOC, ChlaC, TKN-F, TKN-U et SRP-F se retrouvent dans cette liste et sont accompagnées d'une description, respectivement ; Potential of hydrogen, Dissolved organic carbon, Total chlorophyll *a* (corrected for phaeophytin), Total Kjeldahl nitrogen (filtered), Total Kjeldahl nitrogen (unfiltered) et Soluble reactive phosphorus (filtered).

Collaborateurs à la CDD

Une liste des collaborateurs à la CDD a été créée. Elle regroupe les noms de toutes les personnes qui ont contribué de près ou de loin à la récolte des données ainsi qu'à leurs traitements. S'ajoutent également les informations de titres, d'affiliations et de contact pour chaque collaborateur. Cette table de référence nommée « PEOPLE » contient présentement 13 enregistrements et elle est sujette à prendre de l'ampleur dans un futur rapproché. Pour cette raison et dans le même esprit que la librairie de laboratoires, l'abréviation « Lib_ » n'a pas été ajoutée à son nom.

Affiliations

Les affiliations pour chaque collaborateur et laboratoire d'analyse ont été répertoriées dans la table de référence « AFFILIATION ». Elle regroupe actuellement 17 différentes affiliations universitaires et de groupes de recherche. Leur abréviation ou acronyme ainsi que leur adresse y sont listés. Jusqu'à maintenant, les affiliations répertoriées proviennent de 15 universités différentes. Cette librairie est sujette à des modifications fréquentes et c'est pourquoi le nom de la table ne contient pas l'abréviation « Lib_ ».

4.3.3. Librairies taxonomiques

Le développement d'une librairie taxonomique de diatomées est l'étape qui a nécessité le plus d'investissement parmi toutes les librairies créées dans ce projet. La complexité des informations qu'elle contient et leur importance par rapport aux résultats paléolimnologiques explique l'attention qui y a été portée.

La facilité d'identification taxonomique des restes diatomifères trouvés dans une carotte de sédiments dépend de l'état de préservation des spécimens identifiables et de la littérature taxonomique existante. Dans chaque échantillon étudié, une multitude de possibilités de nomenclatures taxonomiques peut rendre l'identification difficile. Les nomenclatures peuvent ainsi référer à différent niveaux de l'échelle taxonomique, c'est-à-dire au niveau du genre, de l'espèce, de la sous-espèce (variété) et de la forme, entre autres. Souvent, lorsque l'identification au niveau de l'espèce est impossible, la méthode privilégiée consiste à

trouver l'équivalent morphologique le plus près, ou encore à donner un nom spécifique au spécimen, lorsque les critères taxonomiques le permettent.

D'un chercheur à l'autre, des différences dans les abréviations, les acronymes et même parfois des fautes de frappe dans les noms latins peuvent être constatés. Ces erreurs ou différences doivent être corrigées et uniformisées afin de rendre les résultats comparables et pour éviter les doublons. Autrement, la crédibilité scientifique des résultats de la CDD serait mise en cause et des interprétations fautives pourraient être engendrées.

Avant de créer la librairie taxonomique principale « Lib_DIATOM », plusieurs étapes de vérifications et d'uniformisation ont été effectuées pour chacune des sources taxonomiques. Ultimement, l'un des objectifs de cette librairie est d'incorporer dans une même table les différents noms de codes et d'identifiants issus de bases de données taxonomiques distinctes. Cette présente section offre d'abord une description des sources utilisées, suivie de la méthode employée pour créer cette librairie.

4.3.3.1. Sources des listes taxonomiques et des données de la CDD

NADED List of Taxa

La liste taxonomique de la North American Diatom Ecological Database (NADED) a été instaurée par le musée américain Academy of Natural Sciences in Philadelphia (ANSP). Elle rassemble les noms et codes taxonomiques de diatomées qui ont été utilisés dans plusieurs projets de recherche nord-américains, dont PIRLA, NAWQA, Environmental Monitoring and Assessment Program-Northeast Lakes (EPA EMAP-NE Lakes) et Diatom Paleolimnology Data Cooperative (DPDC). La coopérative DPDC fournit librement cette liste de 6650 enregistrements. On y dénote plusieurs synonymes et quelques doublons. Néanmoins, l'utilisation de cette liste est utile et recommandée pour les chercheurs, et ce dans le but de produire des résultats d'analyse comparables (Sullivan & Charles, 1994). Pour ces raisons et puisque la majorité des données présentement contenues dans la CDD sont localisées en Amérique du Nord, la NADED List of Taxa servira de structure de base pour la création de la table de référence taxonomique principale de la CDD.

CDD v.1.0

La première version de la Circumpolar Diatom Database comportait déjà une liste de noms et de codes taxonomiques. Celle-ci contenait 590 enregistrements, incluant quelques doublons et synonymes. Les codes pour les noms de diatomées étaient composés par les chercheurs selon la méthode DIATCODE (Hartley, 1986; Williams *et al.*, 1988). Cette liste constituait un premier regroupement de listes taxonomiques issues de différents travaux de recherche, couvrant six régions circumpolaires au Canada et en Russie : le Yukon (Pienitz *et al.*, 1995; Pienitz *et al.*, 1997a-b), le Québec nordique et le Labrador (Fallu & Pienitz, 1999; Fallu *et al.*, 2000) et les régions des rivières Pechora, Lena et Taimyr (Duff *et al.*, 1999; Laing *et al.*, 1999a-b; Laing & Smol, 2000).

Pienitz, R.

Cette liste taxonomique de 214 taxa est issue de travaux de recherche effectués aux environs de Yellowknife, dans les Territoires du Nord-Ouest au Canada (Pienitz *et al.*, 1997a-b). Cette liste est associée à des données paléolimnologiques qui ont été nouvellement incorporées dans la CDD au cours de ce projet. Il en est de même pour les six autres sources citées plus bas. Dans la CDD, le code qui a été attribué à cette région est « YEL », pour Yellowknife.

Saulnier-Talbot, É.

Les données taxonomiques issues des travaux de recherche doctorale de Saulnier-Talbot (2007; Saulnier-Talbot *et al.*, 2007) comptent 265 taxa. Ces données proviennent des villages de Salluit, Kangiqsujuaq et Quaqtaq, situés dans la péninsule d'Ungava, au Nunavik. Dans la nouvelle CDD, le code qui a été attribué à cette région est « SAL », pour Salluit.

Laperrière, L.

Les données des travaux de maîtrise de Laperrière (2007; Rolland *et al.*, 2008; 2009) ne contenaient pas de liste taxonomique distincte, mais elles référaient plutôt aux codes de la liste *CDD v.1*. Ces données proviennent de Southampton Island, dans le sud-ouest du Bassin de Foxe, Nunavut. Le code « SI » a été attribué pour cette région.

Antoniades, D.

Cette liste taxonomique contient 250 taxa et réfère à trois régions de l'archipel arctique canadien qui ont été étudiées par Antoniades *et al.* (2003a-b; 2004; 2005). Ces régions sont : Isachsen sur l'île Ellef Ringnes, Mould Bay sur l'île Prince Patrick et Alert sur les îles Queen Elizabeth, Nunavut. Les codes attribués à ces régions sont respectivement « ISACH », « MB » et « ALERT ».

Côté, G.

La liste taxonomique de Côté *et al.* (2010) contient 84 taxa et elle est issue d'un projet de maîtrise qui s'est réalisé sur l'île Bylot, Nunavut. Le code qui a été attribué à cette région est « BYL ».

Rosén, P.

Les données de Rosén *et al.* (2000) constituent l'une des contributions les plus récentes pour la nouvelle CDD. Réalisés en Suède septentrionale, ces travaux ont entre autres permis d'enrichir la liste taxonomique de la CDD en fournissant une liste étoffée de 755 taxa, comprenant plusieurs informations associées dont des codes de noms de taxa établis selon diverses méthodes. Le code attribué à cette région est « SWE ».

Bigler, C. & Hall, R.

Les données de Bigler & Hall (2002) sont une contribution récente qui a permis d'étendre la couverture de la CDD et d'enrichir son contenu. Elles ont été récoltées au nord de la Suède, dans la région d'Abisko. Le code associé à cette région est « ABK ». La liste taxonomique recueillie contient 292 taxa.

4.3.3.2. Méthode de création de la librairie taxonomique

Pour créer la librairie taxonomique de la CDD qui vise à regrouper les informations de chaque source de données, plusieurs étapes préalables de validation et de réorganisation de données ont été effectuées. Chaque liste a donc été soumise à une réorganisation des données dans le but de rendre leurs informations uniformes et comparables. Cette étape de réorganisation, qui s'est essentiellement effectuée dans MS Excel, consistait, entre autres, à subdiviser les noms de taxa par niveau taxonomique à l'aide de requêtes de déconcaténations. C'est-à-dire que les chaînes de caractères (noms de taxa complets) ont été divisées en plusieurs colonnes dans le tableur, selon un critère de division basé sur les espaces vides. La Figure 13 illustre cette étape de déconcaténation par un exemple détaillé.

Naded ID	Naded Code	Full Taxon name							
52086	Plbasucf	Pinnularia cf. braunii var. amphicephala fo. subconica Venkataraman							
Genre	1er nom Taxa	Espèce	2e nom Taxa	3e nom Taxa	Variété	4e nom Taxa	Forme	Autorité	Year
Pinnularia	cf.	braunii		var.	amphicephala	fo.	subconica	Venkataraman	

Figure 13 : Exemple d'une déconcaténation de nom de taxon.

Au cours de ces étapes préalables au regroupement des données taxonomiques, de nombreuses anomalies ont été dévoilées dans chacune des listes. Parmi celles-ci, on compte des inconstances dans les acronymes et abréviations, des espaces en trop, des points manquants, des doublons, des erreurs de codes, des fautes d'orthographe, etc. Ces anomalies ont été corrigées une à une, ce qui a nécessité de fréquentes vérifications dans la littérature, auprès d'auteurs ou de la communauté scientifique. Une routine de validation a été développée, utilisant des fonctions de recherche automatisées de MS Excel et de l'extension ASAP Utilities. Chaque modification de nature taxonomique apportée aux listes a été notée en historique de modifications, et la valeur d'origine a toujours été conservée à même la liste taxonomique finale. De cette façon, il est possible de connaître l'origine et les raisons des changements apportés.

Une fois les listes taxonomiques adéquatement réorganisées et corrigées, leur importation a été effectuée dans MS Access, plus précisément dans le fichier nommé

« CDD_Taxa.Lib.mdb ». Ce fichier, qui contient toutes les listes taxonomiques réorganisées, a été créé pour servir de référence pour la création de « Lib_DIATOM », cette dernière étant contenue dans la base de données principale (CDD.mdb). Par le biais des fonctions de liaisons de tables entre ces deux fichiers, des requêtes de concordance de noms taxonomiques ont pu être réalisées entre les listes. La Figure 14 est une représentation des relations entre les fichiers physiques « CDD.mdb », « CDD_Taxa.Lib » et leurs listes taxonomiques associées. Les requêtes de concordances ont permis de faire ressortir les taxa absents de la base de données principale qui, rappelons-le, est basée sur la NADED List of Taxa réorganisée et corrigée. Ainsi, si un nom de taxon d'une liste de « CDD_Taxa.Lib » est identique à un nom de « Lib_DIATOM », celui-ci n'est pas ajouté. En contrepartie, si un nom de taxon ne trouve pas de concordance dans la liste principale, la requête affiche ce nom et la décision de l'ajouter ou non peut ensuite être établie. Un index de référence taxonomique a été créé dans le fichier « CDD.mdb », au nom de « INDEX_Taxa.Lib ». Cet index est en fait une table de concordance entre les listes taxonomiques de références et la librairie principale « Lib_DIATOM ». Dans cet index sont répertoriés tous les identifiants uniques des taxa importés, de sorte qu'il est possible de visualiser les noms de taxa, soit dans « Lib_DIATOM » ou encore dans la liste source située dans « CDD_Taxa.Lib ».

Suite à ces étapes de réorganisation et de validation, la liste taxonomique « Lib_DIATOM » créée pour la CDD contient maintenant 7303 enregistrements. Considérant une cinquantaine de retraits de doublons dans les listes originales, cette nouvelle liste présente un gain de quelques 700 taxa par rapport à la liste de référence NADED. Au total, 2472 taxa sont répertoriées dans l'index des tables liées.



Figure 14 : Schématisation des interactions entre les listes taxonomiques de « CDD.mdb » et « CDD_Taxa.Lib ».

4.4. Importation des données

La structure de la base de données une fois finalisée et les librairies créées, l'étape d'importation des données dans les tables MS Access peut être entamée. Avec la même rigueur et le souci d'exactitude appliqué à l'étape de réorganisation et de correction des listes taxonomiques, les données sources à importer ont toutes été vérifiées et ajustées au cours des nombreuses étapes de standardisation préalables. Des protocoles d'uniformisation et de disposition des données ont été établis et sont décrits dans les sections suivantes.

4.4.1. Uniformisation des unités

Avant l'importation des données, leur structure, leurs unités et leur nomenclature ont été uniformisées. Des champs ont été ajoutés pour l'identification des niveaux de carottage, des codes d'échantillonnage, du lac d'origine et du site d'étude. Les données ont été soumises à des tests de validation et certaines ont nécessité des corrections en concertation avec les auteurs. Plusieurs valeurs ont été converties pour être conforme aux unités attribuées aux champs de résultats de la CDD. Les identifiants uniques de « Lib_ANALIT »

correspondants ont été associés à chaque résultat avant leur importation. L'abondance relative pour les résultats d'occurrences diatomifères est l'unité qui est privilégiée pour l'insertion des données. Lorsque les abondances brutes étaient disponibles, seul le total du comptage par niveau a été inséré. De cette façon, il est possible d'obtenir les valeurs brutes par calcul avec le pourcentage. Pour les données d'abondances relatives qui présentaient des décomptes inférieurs à 100% pour les totaux et qui n'ont pas pu être ajustées adéquatement, l'attribution du champ d'analyse « INDET_&_<1% » a été effectuée de même qu'un code de diatomée indéterminé. La définition de ce type d'analyse s'exprime comme suit dans les métadonnées de la table « Lib_ANALIT » : Deduction of the missing percentage, mix of indeterminata, unknown and <1% of abundance species. De cette façon, les erreurs de pourcentages sont assignées à une catégorie distincte, ce qui est une méthode acceptable et usuelle pour les résultats de comptages de diatomées fossiles.

4.4.2. Disposition des tableurs

Les données issues des travaux d'analyse de carotte longue, comme les résultats d'occurrences et des résultats de perte-au-feu (PAF; Loss On Ignition, LOI), sont généralement disposées en plusieurs colonnes dans les tableurs d'origine, et elles peuvent même être réparties sur plusieurs pages. Les données, ainsi disposées, ne peuvent pas être importées telles quelles dans les tables de la base de données relationnelle qui classe les données par lignes plutôt que par colonnes. De plus, l'analyse de carotte longue génère de nombreuses données. Pour ces raisons, leur importation est fastidieuse et requiert de longues manipulations pouvant représenter plus d'une journée de travail par source de données. Le développement d'une méthode de transposition automatique s'est donc avéré nécessaire.

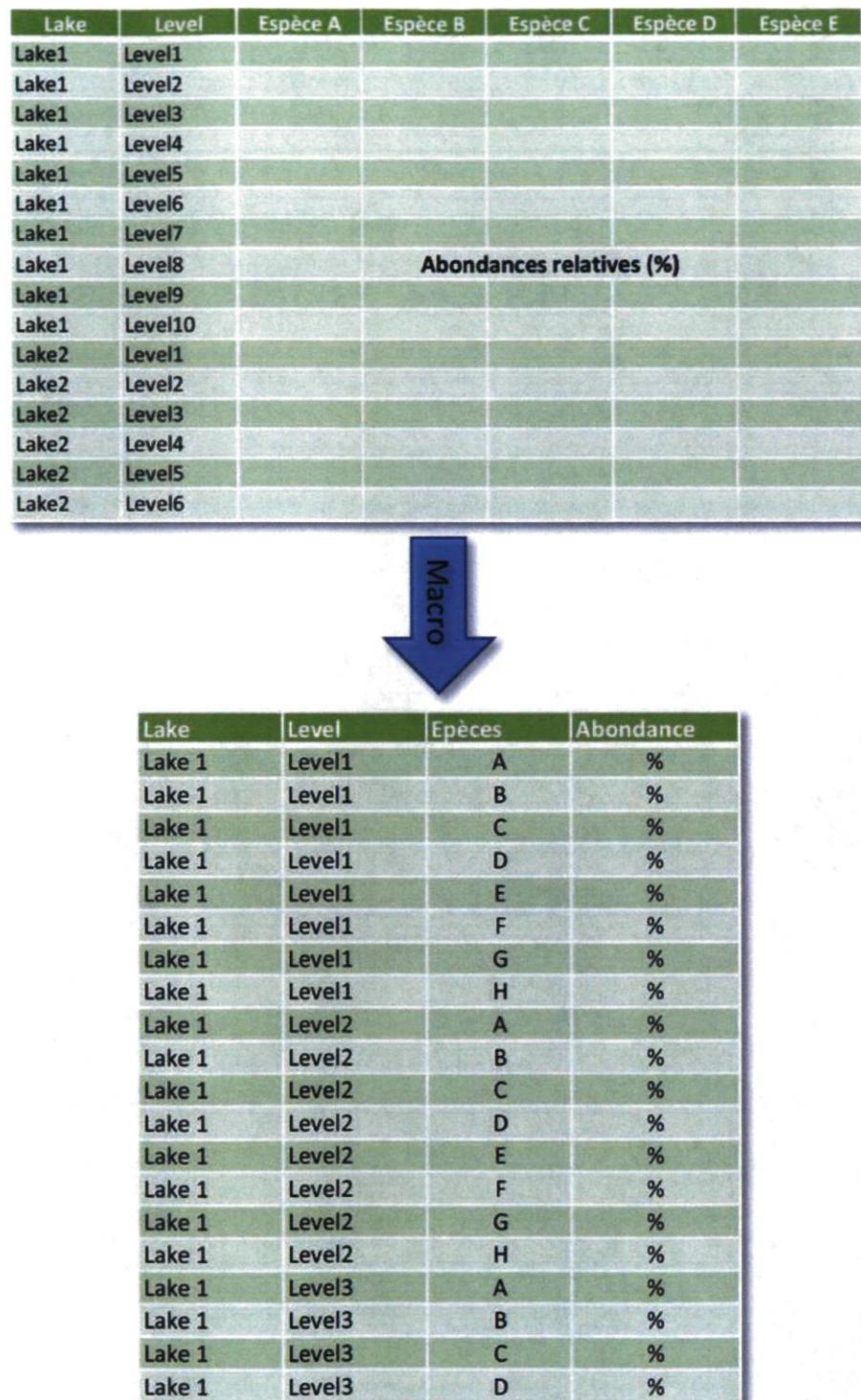
Une macro MS Excel a été créée, ce qui a permis d'automatiser la majorité des étapes de disposition des données avant l'importation. Cette programmation macro, présentée à l'Annexe D, permet de réorganiser sous forme d'enregistrements (lignes) les données autrement présentées sous forme de colonnes. La Figure 15 montre un exemple de ce type de transposition. Cette macro consiste à attribuer les bons codes de lac et de niveau à

chaque résultat, tout en conservant la valeur d'analyse et, dans le cas de données d'abondance, l'identification taxonomique.

Une fois les données ainsi disposées, vérifiées et corrigées, l'importation peut alors être effectuée à l'aide de quelques requêtes SQL dans MS Access. Des modules de programmation VBA⁸ ont été développés dans MS Access afin d'effectuer les correspondances taxonomiques de façon semi-automatisée. Ces modules VBA sont présentés aux annexes C et E⁹. Toutes ces étapes ont été répétées pour chaque source et chaque type de données. Seules les données de localisation, des sites d'étude, de collaborateurs, de publications et d'affiliations ont été insérées de façon manuelle ou par requêtes SQL simples et complexes. Plusieurs corrections et conversions ont été effectuées dans les valeurs de coordonnées géographiques. Les unités privilégiées pour les valeurs X et Y sont les degrés décimaux.

⁸VBA est un langage de programmation Microsoft Visual Basic pour Applications (Visual Basic for Applications).

⁹ Les annexes C et E contiennent également les codes de programmation VBA développés pour les formulaires de consultation MS Access, les sous-formulaires et les modules de recherche.



The diagram illustrates the process of transposing data from a wide format to a long format using a macro in Microsoft Excel.

Wide Format Data:

Lake	Level	Espèce A	Espèce B	Espèce C	Espèce D	Espèce E
Lake1	Level1					
Lake1	Level2					
Lake1	Level3					
Lake1	Level4					
Lake1	Level5					
Lake1	Level6					
Lake1	Level7					
Lake1	Level8					Abondances relatives (%)
Lake1	Level9					
Lake1	Level10					
Lake2	Level1					
Lake2	Level2					
Lake2	Level3					
Lake2	Level4					
Lake2	Level5					
Lake2	Level6					

Long Format Data (Result of Transpose):

Lake	Level	Epèces	Abondance
Lake 1	Level1	A	%
Lake 1	Level1	B	%
Lake 1	Level1	C	%
Lake 1	Level1	D	%
Lake 1	Level1	E	%
Lake 1	Level1	F	%
Lake 1	Level1	G	%
Lake 1	Level1	H	%
Lake 1	Level2	A	%
Lake 1	Level2	B	%
Lake 1	Level2	C	%
Lake 1	Level2	D	%
Lake 1	Level2	E	%
Lake 1	Level2	F	%
Lake 1	Level2	G	%
Lake 1	Level2	H	%
Lake 1	Level3	A	%
Lake 1	Level3	B	%
Lake 1	Level3	C	%
Lake 1	Level3	D	%

A large blue arrow labeled "Macro" points downwards from the wide format table to the long format table, indicating the transformation process.

Figure 15 : Transposition des données par niveau à l'aide d'une macro MS Excel.

5. Chapitre 5 : Résultats

Ce chapitre vise à décrire l'ensemble du nouveau système d'information suite à l'insertion des données et la finalisation des formulaires de consultation. D'abord, une description des fichiers physiques et de leur fonctionnement a été dressée. Ensuite, un dénombrement sommaire du contenu des tables de la base de données est illustré, suivi d'une description de chacun des formulaires de consultation. À la fin du chapitre, les résultats des essais d'intégration de la CDD dans un logiciel SIG sont présentés.

5.1. Fichiers physiques de la CDD

Le nouveau SI développé pour la CDD a généré trois différents fichiers physiques au format « .mdb ». Un fichier principal, aussi le plus volumineux, contient l'ensemble des tables, leurs relations et les données de la CDD. Ce fichier se nomme « CDD.mdb » et il est d'une taille de 7,56 mégaoctets (Mo). À ce fichier principal s'ajoutent deux autres fichiers qui y sont liés (Figure 16). D'une taille de 0,99 Mo, le fichier « CDD_Taxa.mdb » contient les différentes listes taxonomiques originales sur lesquelles nous nous sommes basés pour constituer la librairie « Lib_DIATOM ». Cette librairie se retrouve dans le fichier « CDD.mdb », mais dans le cas d'une mise à jour ou d'une vérification d'informations de taxon, elle pourra être connectée aux listes sources de « CDD_Taxa.mdb ». En effet, une fonction de liaison de tables externes permet de communiquer entre ces deux fichiers. Cette liaison a été créée avec des restrictions de modification du fichier principal qui contient les données. Cela a pour but de préserver l'intégrité des données de la BD qui pourrait être perturbée par des mauvaises manipulations ou de possibles erreurs logicielles. Il en est de même pour la liaison avec le troisième fichier « CDD_Home.mdb ». D'une taille de 4,65 Mo, ce fichier contient les formulaires de consultation qui ont été développés dans le cadre de ce projet. Ces formulaires permettent de naviguer dans le contenu de « CDD.mdb » par liaison. Ici encore, il s'agit d'une façon sécuritaire de consulter les données du fichier principal. En plus des raisons de sécurité, cette structure physique en trois fichiers a été conçue afin de faciliter l'éventuelle diffusion de la CDD parmi la communauté scientifique. En effet, dans l'éventualité où le fichier principal « CDD.mdb » serait instauré sur un serveur réseau, seul le fichier de formulaires « CDD_Home.mdb »

devra être distribué aux utilisateurs. Ce détail simplifie la gestion du SI, des autorisations d'accès et d'éventuelles mises à jour de données.

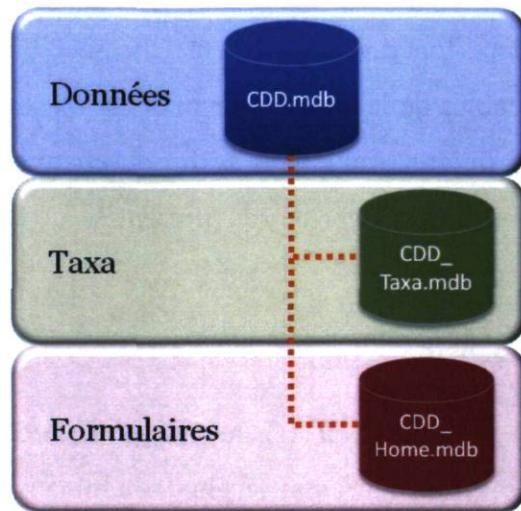


Figure 16 : Représentation des fichiers physiques de la CDD.

5.2. Sommaire du contenu de la CDD

Grâce aux contributions de la communauté scientifique et certaines recherches récentes du LPA, la nouvelle version de la CDD a plus que triplé son volume d'informations. Elle renferme maintenant quelques 572 sites d'échantillonnage (vs 261 dans la première version), 39 350 données d'occurrence (vs 10 072) et 15 274 résultats d'analyse d'eau (vs 3 366). Sa couverture s'est accrue et regroupe maintenant 18 régions circumpolaires (vs 6) réparties sur trois continents : Amérique du Nord, Europe et Asie (Sibérie). Le Tableau 2 résume en trois catégories la quantité d'informations que contient maintenant la CDD.

Tableau 2 : Sommaire du contenu de la CDD (novembre 2010).

<i>Les Données :</i>	
18	Régions circumpolaires arctiques
15	Expéditions de terrain
572	Lacs
1 117	Échantillonnages
549	Carottes de sédiments
1 343	Niveaux
39 350	Résultats d'occurrences
1 788	Résultats LOI
568	Échantillons d'eau
15 274	Résultats d'analyse d'eau
<i>Les Métadonnées :</i>	
6	Laboratoires d'analyse
9	Collaborateurs, responsables de projets
17	Affiliations
20	Publications associées aux données
<i>Les Librairies :</i>	
7 303	Taxon de diatomées
1169	Taxon avec résultats d'occurrence
92	Types d'analyses
7	Catégories d'analyse
234	Unités
246	Pays
3039	Universités

5.3. Description et fonctionnement des formulaires MS Access

CDD Home page

À l'ouverture du fichier de consultation de la CDD, soit « CDD_Home.mdb », un formulaire de menu principal apparaît automatiquement. Il s'agit du formulaire intitulé « CDD Home page », celui-ci est illustré à la Figure 17. Le nom de base de données et les logos du LPA, du CEN et de l'Université Laval y figurent en premier plan. Un aperçu de la couverture spatiale de la CDD est visible en arrière-plan. Quatre options de navigation sont proposées aux utilisateurs par des touches interactives. Les deux premières options permettent d'effectuer une recherche de résultats soit par lac (Search by lake), soit par taxon (Search by taxon). Les deux autres options permettent de consulter les publications associées aux données de la CDD (Publications) ou les collaborateurs de la CDD (People).

Lorsque l'utilisateur actionne l'une de ces touches de navigation, le formulaire correspondant s'ouvre en premier plan devant le formulaire de menu principal. Ce dernier demeure toujours en arrière-plan, à moins que l'utilisateur décide de fermer cette fenêtre. Si l'utilisateur désire ouvrir à nouveau la fenêtre de menu principal qui a été fermée, la barre de menu latérale (Figure 18) permet de réaliser cette action à tout moment. Il en est de même pour tous les autres formulaires et sous-formulaires disponibles dans cette application.

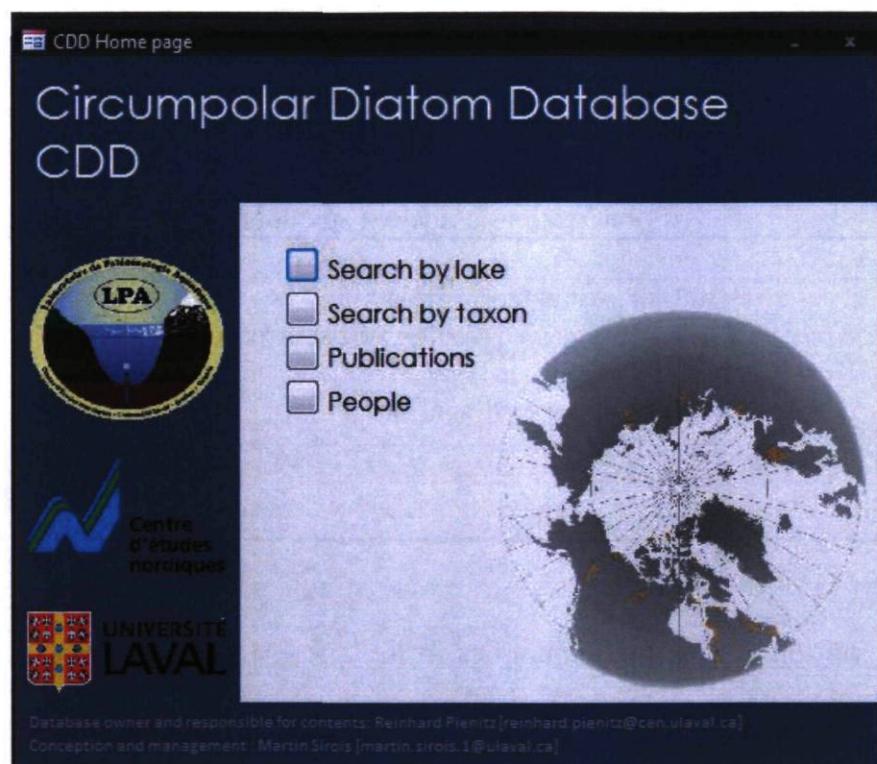


Figure 17 : Formulaire de menu principal « CDD Home page ».



Figure 18 : Extrait de la barre de menu latérale de CDD_Home.mdb.

Results by lake

Lorsque la touche de recherche par lac (Search by lake) est actionnée soit à partir du menu principal ou de la barre de menu latérale, le formulaire intitulé « Results by lake » s'ouvre en premier plan. Ce formulaire permet d'obtenir plusieurs types d'informations à partir d'une recherche basée sur les lacs. Dans l'entête du formulaire, une série d'options de recherche s'aligne horizontalement et se termine par un bouton interactif de recherche (Figure 19). Ce module a été conçu afin que l'utilisateur puisse sélectionner différents critères lui permettant d'atteindre un lac ciblé. Il permet de filtrer les lacs contenus dans la CDD d'abord par pays, puis, soit par État ou par chercheur. Lorsque ces premières étapes de filtrage sont appliquées, l'encadré « Lake » à la droite de l'entête fournit une liste des codes de lac résultant de cette recherche. Les codes de lac affichés dans cette liste déroulante peuvent également être accompagnés des noms de lac, lorsque ceux-ci sont disponibles. Une fois le lac sélectionné, l'utilisateur peut appliquer la recherche de résultats pour celui-ci. Plusieurs informations s'afficheront alors dans le formulaire et ses sous-formulaires (les onglets).

Dans le coin supérieur gauche se trouvent des informations de type géographique et physique, soit les coordonnées en degrés décimaux, la profondeur, la circonférence et la superficie du lac en question.

Au centre de la partie supérieure du formulaire, deux onglets fournissent des informations concernant l'échantillonnage de la carotte de sédiments ainsi que des prélèvements d'eau. Dans le premier onglet (celui de gauche), cinq informations sont affichées à propos de la carotte de sédiments, soit le code de la carotte, sa longueur totale, les dates de début et de fin des travaux de terrain pour l'ensemble du projet de recherche, ainsi que la date précise de l'échantillonnage, lorsque disponible. Les informations fournies dans l'onglet d'échantillon d'eau sont les mêmes, à l'exception qu'il n'y a pas de longueur totale et que le code d'échantillon est remplacé par l'identifiant unique (ID) de l'échantillon d'eau.

À la droite des onglets relatifs aux échantillonnages se trouve une liste des publications correspondantes au projet de recherche pour le lac sélectionné. Cette liste possède des liens

réactifs, de sorte que l'utilisateur peut ouvrir le formulaire de publications (Publications) par un clic de souris sur la publication désirée. Ce formulaire est décrit plus bas.

La partie inférieure de ce formulaire présente quatre onglets de résultats. Le premier qui est affiché par défaut est celui des abondances relatives de diatomées (Diatom List) sous forme de liste. Ce sous-formulaire permet de consulter les résultats d'occurrence en pourcentage. Les lignes représentent les niveaux de carottage, tandis que les colonnes représentent les taxa de diatomées. Un menu déroulant de taxa est accessible à la case « Full_Taxon_Name » et permet de sélectionner le ou les taxa désirés pour la consultation. Le nom de chaque taxon choisi s'affiche au haut de la colonne correspondante. Ce sous-formulaire utilise une requête de type tableau croisé dynamique. Le tableau original qui regroupe tous les résultats de tous les lacs présentant des résultats d'occurrences est accessible via la barre de menu latérale au nom de « Abund. (%) PivotTable ».

Full_Taxon_Name			
Achnanthes curtissima Carter	Achnanthes kueibsi Lange-Bertalot	Achnanthes levanderi Hustedt	
0	0,5	2,280130293	
0,5	1	1,967213115	0,655737705
1	1,5	1,639344262	1,967213115
1,5	2	0,931677019	1,863354037
2	2,5	2,597402597	0,324675325
2,5	3	1,261829653	
3	3,5	0,323624595	3,236245955
3,5	4	0,986842105	1,644736842
4	4,5	0,660066007	0,99009901
4,5	5		0,628930818
5	5,5	0,326797386	1,633986928
5,5	6	0,328947368	0,328947368
6	6,5		0,323624595
6,5	7	0,322580645	

Figure 19 : Formulaire de recherche de résultats par lac « Results by lake ».

Le deuxième onglet, intitulé « Diatom Graph », fournit également des résultats d'occurrences diatomifères (Figure 20). Ces résultats sont présentés sous forme graphique, soit le mode de représentation le plus commun en paléolimnologie. Il s'agit d'un histogramme à bandes horizontales où l'axe des ordonnées inversé représente la profondeur dans la carotte et l'axe des abscisses les abondances relatives pour une diatomée donnée. Encore ici, il est possible de sélectionner les taxons voulus à l'aide d'un menu déroulant. Ce sous-formulaire emploie une requête de graphique croisé dynamique dont la version complète est accessible via la barre de menu latérale au nom de « Abund. (%) Graph ».

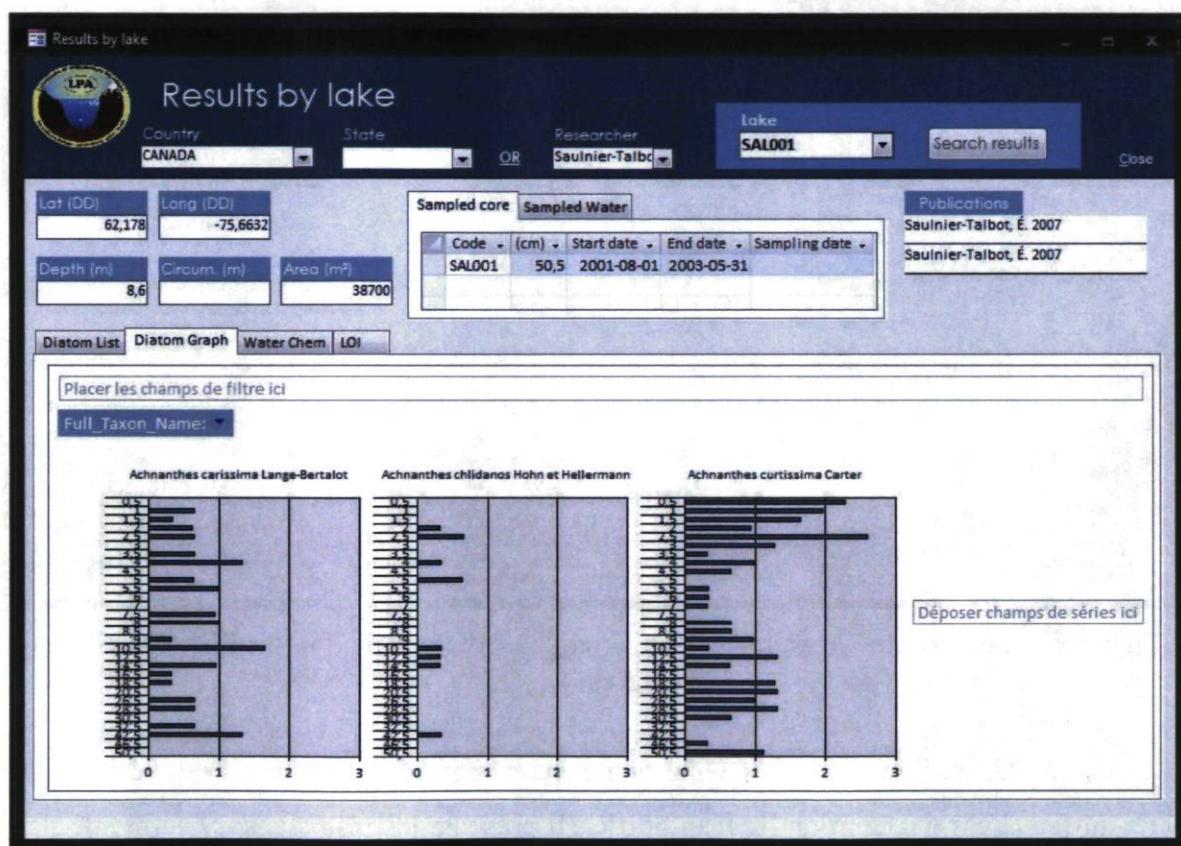


Figure 20 : Formulaire « Results by lake », onglet « Diatom Graph » (graphiques d'abondances relatives).

Le troisième onglet (Figure 21) proposé dans ce formulaire réfère aux résultats d'analyse d'eau *in situ* et *ex situ*. Cet onglet nommé « Water Chem » offre une liste de noms d'analyses, les résultats associés, le symbole unitaire de chaque résultat ainsi que la date d'échantillonnage lorsque disponible. Dans ce sous-formulaire, des options de tri et de filtrage sont offertes aux utilisateurs pour chacune des colonnes. Il est également possible

d'effectuer une recherche parmi les résultats affichés, à l'aide de la case « Rechercher » au bas de la liste. Tout près de cette case, des flèches de navigations permettent d'explorer les enregistrements de cette liste, bien que l'utilisateur puisse tout autant employer la barre de défilement verticale qui est située à l'extrême droite.

Name	Analyte	Result	Unit term	Date Sampling
Ca		13,5	mg/L	
CHLaC		0,5	µg/L	
CHLaU		0,8	µg/L	
Cl		4,8	mg/L	
COND		100	µS/cm	
DIC		12,1	mg/L	
DIST		120	km	
DOC		5,6	mg/L	
Fe		36,6	µg/L	
K		1,95	mg/L	
Mn		4	µg/L	
Na		4,66	mg/L	
NH3		ADL	µg/L	
NO2		2	µg/L	
NO3		ADL	µg/L	
pH		8,5	pH	
PN		60	µg/L	
POC		488	µg/L	

Figure 21 : Formulaire « Results by lake », onglet « Water Chem » (résultats d'analyses d'eau).

À la Figure 22 est illustré le quatrième onglet de ce formulaire, il s'agit des résultats de perte-au-feu (Loss On Ignition; LOI) à 550°C. Il présente les résultats sous forme de liste par niveaux de carottage. Au même titre que l'onglet « Diatom List » précédemment décrit, il est issu d'une requête de tableau croisé dynamique. Dans cette liste, chaque ligne représente un niveau et les colonnes réfèrent aux résultats par type d'analyse LOI. Les résultats de l'analyse LOI-550 sont exprimés en pourcentage du poids sec de la matière organique totale.

Start_Depth	End_Depth	Result
0	0,5	15,3351612
0,5	1	15,70234114
1	1,5	15,59439379
1,5	2	16,0728855
2	2,5	15,6807246
2,5	3	15,75947563
3	3,5	16,49484536
3,5	4	16,32842758
4	4,5	15,5272825
4,5	5	14,81442093
5	5,5	14,7479776
5,5	6	14,30398426
6	6	14,50000000

Figure 22 : Formulaire « Results by lake », onglet « LOI » (perte-au-feu).

Results by Diatom

À partir du menu principal ou de la barre de menu latérale, il est possible de sélectionner l’option de recherche par taxon de diatomées (Search by taxon). Lorsque cette option est choisie, le formulaire « Results by Diatom » s’ouvre en premier plan (Figure 23). L’entête de ce formulaire offre des options de recherche étoffées par le biais de cinq cases de menu déroulant. Ce module de recherche fonctionne de façon progressive selon la hiérarchie des niveaux taxonomiques. Il offre la possibilité de chercher en fonction des niveaux taxonomiques élémentaires tout en permettant de trouver les noms plus complexes de taxa, comme dans le cas de taxa spécifiques à un projet de recherche particulier. Plus précisément, la première lignée de cases de recherche permet de filtrer les noms de taxa par rangs taxonomiques de premier et de deuxième niveau, sans toutefois exclure les noms qui présentent des caractères spéciaux à travers ces mêmes rangs. La liste de noms résultant de ce filtrage est exposée dans la case inférieure intitulée « Taxon search match ». Lors d’une

recherche de taxon, la première étape consiste à sélectionner un genre diatomifère. Cette sélection affichera par la suite une liste de choix dans la case subséquente « Species ». Déjà, la case de résultats de noms de taxa affichera une liste des taxa pour lesquels il est possible d'obtenir des résultats d'analyse. L'utilisateur peut décider de chercher le taxon ciblé dans cette liste, ou encore raffiner cette première liste de noms en sélectionnant davantage de critères de recherche. Pour raffiner la liste de noms, il s'agit de poursuivre la sélection dans les choix proposés par la case « Species ». Ici, comme dans les cases suivantes « Third Taxon name, variety » et « Fourth Taxon name, form », il est possible de sélectionner un nom précis ou encore de sélectionner la première ligne de ces menus. La première ligne vide de ces cases, lorsque sélectionnée, permet d'obtenir tous les noms de taxa peu importe si leur nom de rang est manquant ou irrégulier. Cela permet donc d'accéder aux taxa spécifiques à un projet de recherche en particulier (ex. *Achnanthes* sp. 1 EMILIE), autrement exclus des choix de taxon par filtrage excessif.

Une fois que le taxon ciblé a été trouvé dans la liste de choix « Taxon search match », l'utilisateur peut obtenir les résultats correspondant à ce taxon en appuyant sur la touche interactive « Results for this taxon ». La grille située dans la partie inférieure du formulaire affichera la liste des sites où le taxon a été observé. Le code du lac, son nom, lorsque disponible, le niveau dans la carotte où le taxon a été observé, ainsi que l'abondance relative du taxon s'y retrouvent listés. Cette liste offre les mêmes options de tri, de filtre, de recherche et de navigation que pour l'onglet « Water Chem » du formulaire « Results by lake » précédemment décrit.

À partir des résultats affichés dans cette liste, une opération supplémentaire peut être effectuée par l'utilisateur. En effet, le bouton « See Water Chem for this lake » situé au pied du formulaire permet d'afficher les résultats des analyses d'eau pour un lac sélectionné. Lorsque le bouton est activé par l'utilisateur, le formulaire « Results by lake » s'ouvre à l'onglet « Water Chem » et présente les résultats du lac sélectionné.

Results by diatom

Results by Diatom

Close

Genus	Species	Third Taxon name, variety	Fourth Taxon name, form
Achnanthes	altaica		
Achnanthidium	amoena		
Actinella	austriaca		
Adlaafia	biasolettiana		

Taxon search match

Achnanthes austriaca var. helvetica

Results for this taxon

Lake code	Lake name	Start depth	End depth	Relative Abundance (%)
ABK003		0	1	0,70921985816
ABK005		0	1	0,90293453725
ABK006		0	1	0,2493765586
ABK008		0	1	0,24691358025
ABK009		0	1	0,49261083744
ABK011		0	1	0,95238095238
ABK017		0	1	0,49627791563
ABK018		0	1	4,2755344418

Enr : 1 sur 49 | Rechercher | See Water Chem for this lake

Figure 23 : Formulaire « Results by Diatom ».

Publication details

Le formulaire « Publication details » est accessible depuis trois endroits, soit dans le menu principal, dans la barre de défilement latérale ou à partir de la case correspondante du formulaire « Results by lake ». Ce formulaire (Figure 24) fournit des informations de nature bibliographique concernant les projets contenus dans la CDD. Le nom complet de la publication est l'information qui figure en premier à l'ouverture de ce formulaire. Les cotes DOI, ISBN et ISSN y sont également affichés comme informations complémentaires. Dans la section du bas, un sous-formulaire fournit une liste des régions auxquelles réfère la publication affichée et pour lesquelles la CDD possède des résultats qui y sont associés. L'utilisateur peut explorer les publications à l'aide des flèches de navigation au pied du formulaire. Une recherche par contenu peut également être effectuée, il est donc facile de

trouver une publication à partir d'une recherche par nom d'auteur ou encore par mot-clé, entre autres.

Au bas du formulaire, un bouton intitulé « First Author details » permet d'obtenir davantage d'informations à propos de l'auteur principal, soit le premier nom qui figure dans le titre complet de la publication. Lorsque l'utilisateur en fait la demande, ce bouton permet d'ouvrir le formulaire des collaborateurs « People » et de positionner ce dernier directement sur la page correspondante au premier auteur d'une publication sélectionnée. Ce formulaire de collaborateurs est décrit plus en détail dans la section suivante.

Dans le cas où le formulaire « Publication details » aurait été ouvert à partir de la case « Publications » du formulaire « Results by lake », seule la publication préalablement sélectionnée sera affichée. Cela est possible grâce à une requête de filtre. Le filtre peut toutefois être désactivé au besoin, en appuyant sur la touche « Filtre » au pied de la fenêtre de publications, jusqu'à ce qu'elle affiche plutôt « Non filtré ».

The screenshot shows a software window titled 'Publication details'. At the top left is a circular logo with 'T.P.A.' and a globe icon. The main title 'Publication details' is centered above a text area. The text area contains the following information:

Publication name
Saulnier-Talbot, E., Leng, M.J. and Pienitz, P. 2007. Recent climate and stable isotopes in modern surface waters of northernmost Ungava Peninsula, Canada. Canadian Journal of Earth Sciences.

Below this are three input fields: DOI (10.1139/E06-089), ISBN (empty), and ISSN (empty). A section titled 'Related field expedition' contains a table with the following data:

Region code	Region name	Start Date
SAL	Salluit, Kangiqsujuaq and Quaqtaq, Ungava	2001-08-01

At the bottom of the form are buttons for 'First Author details', navigation (Enr: 14 1 8 sur 18), filtering ('Non filtré'), and search ('Rechercher').

Figure 24 : Formulaire « Publication details ».

People

Le dernier formulaire de consultation se nomme « People » et il regroupe les informations relatives aux collaborateurs de la CDD. Il est accessible par le menu principal, la barre de menu latérale ainsi qu'à partir du bouton « First Author details » du formulaire « Publication details » précédemment décrit. Son entête comporte un menu déroulant qui permet de sélectionner un collaborateur ciblé. Lorsque déroulé, ce menu présente le nom, prénom et l'université de chaque collaborateur de la CDD sous la forme d'une liste continue. Quatre onglets fournissent différents regroupements d'informations par rapport à la personne sélectionnée dans le menu, dont les informations de contact (Details), le ou les affiliation(s), les territoires de recherche (Related field expedition) et les publications. Entre autres, dans l'onglet « Details » se retrouvent le titre et l'adresse principale du bureau du chercheur (Figure 25). Le deuxième onglet fournit la liste des affiliations auxquelles sont rattachés les chercheurs, tandis que l'onglet des terrains de recherche dresse la liste des

sites d'étude pour lesquels des données ont été ajoutées à la CDD. Un dernier onglet dénombre les publications du chercheur relatives aux contributions pour la CDD (Figure 26).

The screenshot shows a software interface titled 'People' with a logo of a globe and the acronym 'LPA'. The window title is 'People' and the search bar shows 'List Saulnier-Talbot'. The main area has tabs: 'Details' (selected), 'Affiliation', 'Related Field', and 'Publication'. The 'Details' tab displays the following information:

Last name	Saulnier-Talbot
First name	Émilie
Title	Science Coordinator (CEN)
University affiliation	Université Laval
University	Université Laval
Univ. address	2325, rue de l'Université Québec (Québec) G1V 0A6 Canada

The 'Contact details' section contains the following fields:

Address	Pavillon Abitibi-Price, 2405, rue de la Terrasse, Université Laval
Postal code	G1V 0A6
Phone	1 (418) 656-2131 poste 2503
Fax number	[empty]
Email	emilie.saulnier-talbot@cen.ulaval.ca
Webpage	http://www.cen.ulaval.ca/en/membre.aspx?id=1604

At the bottom, there are navigation buttons: 'Enr: 14 < 12 sur 13 > >>' and 'Aucun filtre Rechercher'.

Figure 25 : Formulaire « People » onglet détails.

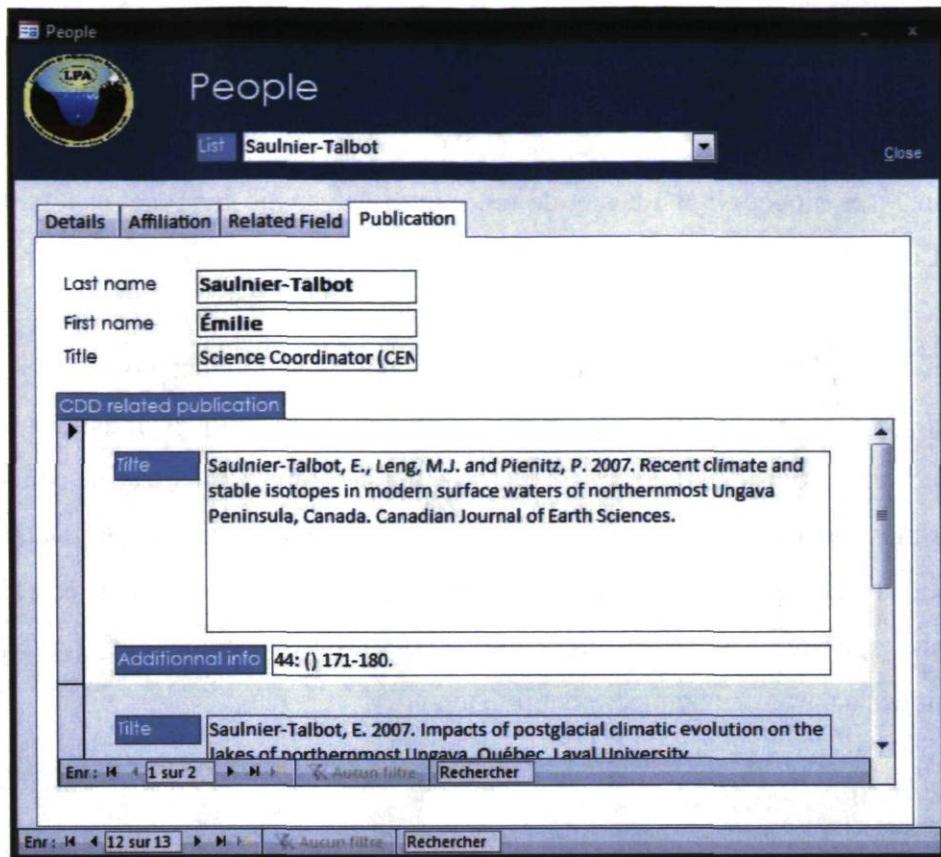


Figure 26 : Formulaire « People », onglet publications.

5.4. Tests d'utilisation dans un logiciel SIG

Le système d'information tel que conçu dans ce projet, étant basé sur le modèle relationnel et le langage SQL, possède un potentiel de connectivité avec plusieurs applications ou logiciels. Parmi ceux-ci, on compte les logiciels de cartographie assistée par ordinateur, communément appelés logiciels SIG. Ces logiciels peuvent utiliser les données issues de divers supports, tels que des tableurs, des bases de données réseau, des bases de données locales ou encore des systèmes d'information géographique.

Dans sa forme actuelle, il est facile d'extraire une grande quantité de résultats de la CDD afin de pouvoir les utiliser dans un logiciel SIG. Pour ce faire, il s'agit d'effectuer une requête dans MS Access et de choisir l'un des nombreux formats de données offerts avant l'extraction. Cependant, par souci de productivité et dans l'optique d'une diffusion réseau

ou publique de la CDD, nous avons voulu tester le potentiel de communication du SI avec un logiciel SIG en excluant l'étape d'exportation. Cette fonctionnalité explorée ici pourrait simplifier l'utilisation des données de la CDD et ultimement contribuer à valoriser les projets d'analyses biogéographiques et de reconstitutions spatio-temporelles à partir de ces résultats. Cette section présente les résultats issus de deux tests de connexion avec le logiciel SIG ArcGIS de ESRI (version 9.3.1). Ce logiciel est couramment utilisé au LPA, au CEN et dans plusieurs grandes institutions à travers le monde. Pour des raisons de compatibilité et d'accessibilité au support et aux ressources techniques, il a été choisi parmi d'autres pour les besoins de cet exercice. Dans ce qui suit, un premier test de connexion est décrit : il s'agit de la connexion OLE DB (Object Linking and Embedding, Database). En second lieu, la description de l'importation de l'ensemble de la CDD directement dans ArcGIS est rapportée. À la fin de cette section, des exemples de représentations cartographiques à partir des données sont illustrés.

Connexion OLE DB

Le logiciel ArcGIS, comme bien d'autres logiciels d'ailleurs, permet l'utilisation d'une base de données à l'aide d'une connexion Object Linking and Embedding, Database (OLE DB). La connexion OLE DB, est une interface de programmation qui permet à un programme d'accéder à diverses sources de données. Cette option de connexion est très répandue, en plus d'être simple d'utilisation. La CDD est physiquement constituée d'un format de fichier « .mdb », celui-ci peut être distribué soit tel quel ou en réseau. Conséquemment, il est fort envisageable d'employer une telle connexion pour la consultation de la CDD dans un SIG, ce qui justifie le présent test d'utilisation.

Pour établir la connexion OLE DB, l'utilisateur doit avoir accès à la lecture du fichier « CDD.mdb ». Pour ce faire, l'utilisateur peut posséder une version physique du fichier ou détenir d'un droit d'accès via une connexion réseau. Le logiciel ArcGIS permet d'établir une connexion par l'une ou l'autre de ces options. Ces dernières sont disponibles dans le module ArcMAP, à l'onglet « File » de la barre de menu horizontale en en-tête, puis à l'option « Add Data... » (ArcGIS version 9.3.1). La fenêtre « Add Data » s'ouvrira alors, il s'agit de choisir l'option « Database Connections » puis « Add OLE DB Connection » afin

d'établir la connexion. Suite aux tests effectués dans ce projet, les propriétés de liaisons recommandées pour les fournisseurs OLE DB sont Microsoft Jet 4.0 OLE DB Provider, Microsoft OLE DB Provider for ODBC Drivers, ou Microsoft OLE DB Provider for SQL Server, dépendant des propriétés d'accès au fichier. Enfin, l'utilisateur doit choisir l'emplacement de la source de données, c'est-à-dire CDD.mdb.

Une fois la connexion établie, les tables et leurs données de la CDD peuvent être consultées directement dans ArcGIS. Les relations, leurs propriétés et les valeurs sont fidèles à la base de données source. Les données sont donc accessibles et conformes; toutefois, elles ne sont pas automatiquement positionnées dans l'espace. Pour ce faire, une requête de sélection doit être effectuée afin d'extraire une table qui contiendra les données désirées pour la représentation graphique. Cette table se nomme « vue » dans de tels cas, puisqu'il ne s'agit que d'une sélection de données. La requête de sélection peut être réalisée dans ArcGIS, ou encore directement dans MS Access préalablement à la connexion OLE DB. Depuis cette vue résultante, les coordonnées des lacs doivent ensuite être sélectionnées par l'utilisateur, un système de référence spatial doit être choisi ainsi que le type de projection. Cette étape s'effectue à partir de l'option « Add XY Data... » de l'onglet « Tools » dans la barre de menu horizontale.

Une fois les données ponctuelles géoréférencées, la consultation de la base de données à partir de la section graphique du module ArcMAP a été testée. À l'aide de l'outil « Identify » disponible à partir de l'icône «  » de la barre d'outils « Tools », il est possible d'afficher les informations d'un lac sélectionné (correspondant à un point dans la fenêtre principale d'ArcMAP). Lorsqu'appliquée, cette action ouvre une fenêtre qui illustre sous forme de liste tous les champs de vue qui sont issus de la requête de sélection. Il est possible de choisir le champ de tri par défaut qui est affiché dans l'explorateur de cette fenêtre, voir l'encadré rouge de la Figure 27. Nous constatons que, peu importe le champ de tri choisi, la hiérarchie de ceux-ci n'est pas adéquate. Dans le cas de la Figure 27, le champ des niveaux de carottage qui a été choisi par défaut illustre bien cette constatation. En effet, chaque champ qui contient un résultat s'affiche dans l'explorateur, rendant difficile la consultation des résultats.

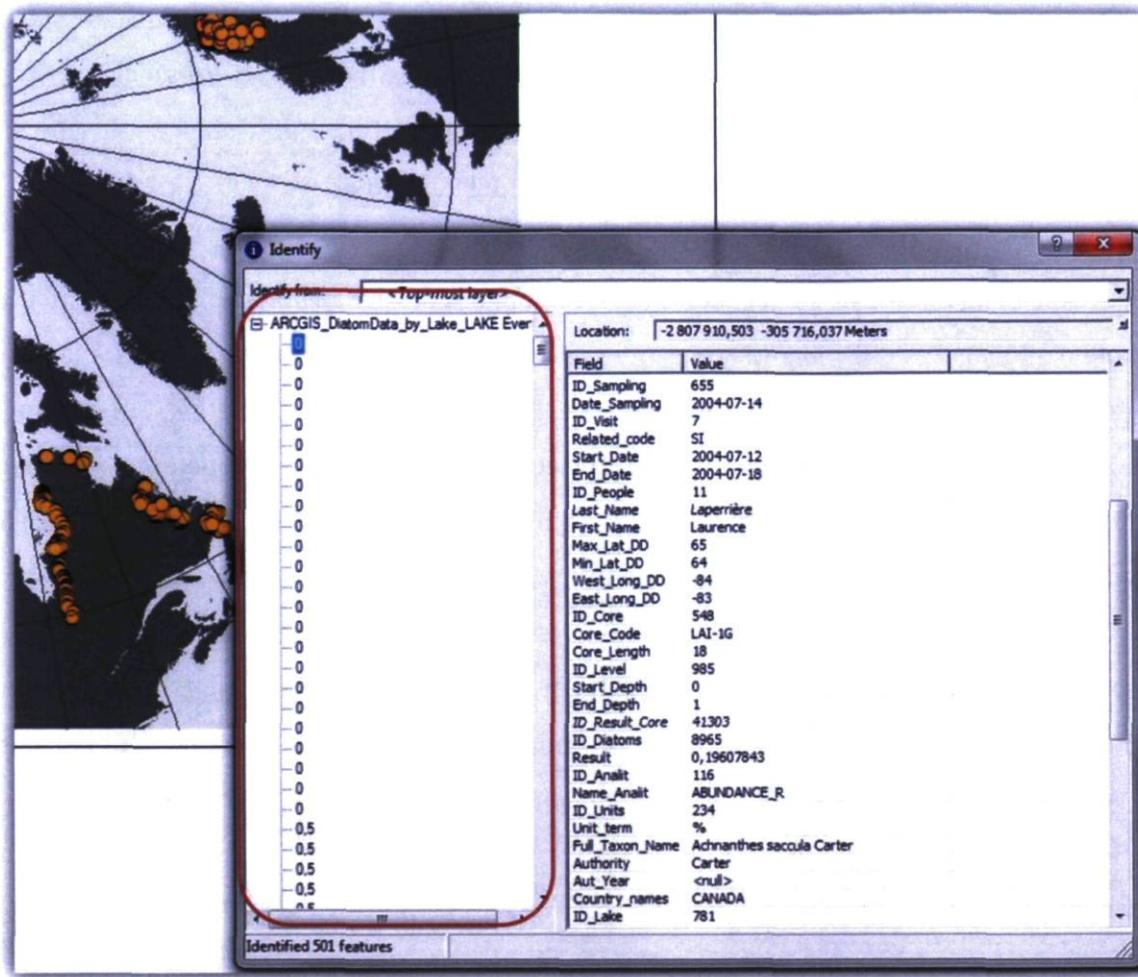


Figure 27 : Détail de l'utilisation de la connexion OLE DB dans ArcGIS.

Concernant les requêtes spatiales, l'utilisation de la CDD par connexion OLE DB ne permet pas la sélection spatiale dans ArcGIS. En effet, la sélection spatiale des objets géoréférencés, qui peut habituellement s'effectuer à l'aide de l'outil « Select by Location », n'est pas applicable dans cette situation.

Des lectures au sujet des connexions OLE DB entre MS Access et ArcGIS permettent d'identifier la cause principale de ces problèmes de consultation et de sélection spatiale (ESRI Support, 2010). Le problème résiderait dans le type d'identifiant unique employé par l'une ou l'autre de ces applications. La base de données MS Access a été créée en utilisant un identifiant unique indépendant d'une table à l'autre. Concrètement, cela signifie pour une table A, que le premier enregistrement créé portera le numéro d'identifiant 1.

Cependant, pour une table B, l'identifiant d'un premier enregistrement peut également être 1. Le SGBD fait par la suite une distinction entre la table, ce qui lui permet de différencier les deux ID. Or, ArcGIS gère différemment les ID, de sorte qu'il exige que chaque identifiant possède un numéro unique pour chaque table d'une même BD. Ce champ, que ESRI nomme « ObjectID » (OID), doit être celui rattaché aux relations dans la BD, soit les clés primaires et secondaires. Sur la base de cette information, nous avons procédé à un second test d'utilisation. La section qui suit décrit les observations obtenues suite à la conversion de la CDD MS Access en Personal Geodatabase ESRI pour son utilisation dans ArcGIS.

Conversion de la CDD en Personal Geodatabase

Ce deuxième essai d'utilisation de la CDD par connectivité directe utilise une toute autre approche que la connexion OLE DB. En effet, il s'agit ici de convertir l'ensemble des tables, des relations et des données de la CDD dans un format de base de données proposé par ArcGIS, soit la Personal Geodatabase (PersonalGDB). Ce format de base de données a été conçu à partir du module ArcCatalogue offert dans la suite ArcGIS. L'une des particularités de ce type de base de données est le fait que les identifiants uniques sont en ObjectId, que les relations créées dans ArcCatalogue ne sont pas visibles dans MS Access, et que la table de lacs contenant les objets géoréférencés possède des attributs géospatiaux (champ nommé « Shape »). La PersonalGDB résultante est un fichier .mdb compatible avec plusieurs SGBD. Toutefois, les modifications apportées à partir d'un SGBD autre qu'ArcGIS corrompront la PersonalGDB qui ne sera plus adéquate pour utilisation dans ArcGIS. Cela signifie que, si cette option est choisie pour la gestion future de la CDD, les gestionnaires devront posséder une version d'ArcGIS afin de maintenir la base de données, ce qui limite l'accès à la CDD dans la communauté.

L'utilisation de la CDD convertie en PersonalGDB dans ArcGIS offre plus de possibilités que lors de l'essai de connexion OLE DB. Premièrement, la sélection spatiale des lacs est désormais rendue possible et toutes les tables, les relations et les requêtes de la CDD sont directement accessibles et modifiables à partir des différents modules d'ArcGIS. De plus, la consultation des données à partir de l'outil « Identify » est plus simple. En effet, la

hiérarchie des relations de la PersonalGDB est maintenue, ce qui fait que la consultation des tables depuis l'explorateur est mieux ordonnée. La Figure 28 illustre bien que l'organisation des données, suite à la sélection des champs par défaut pour chacune des tables, est beaucoup plus claire et intuitive qu'à l'essai précédent.

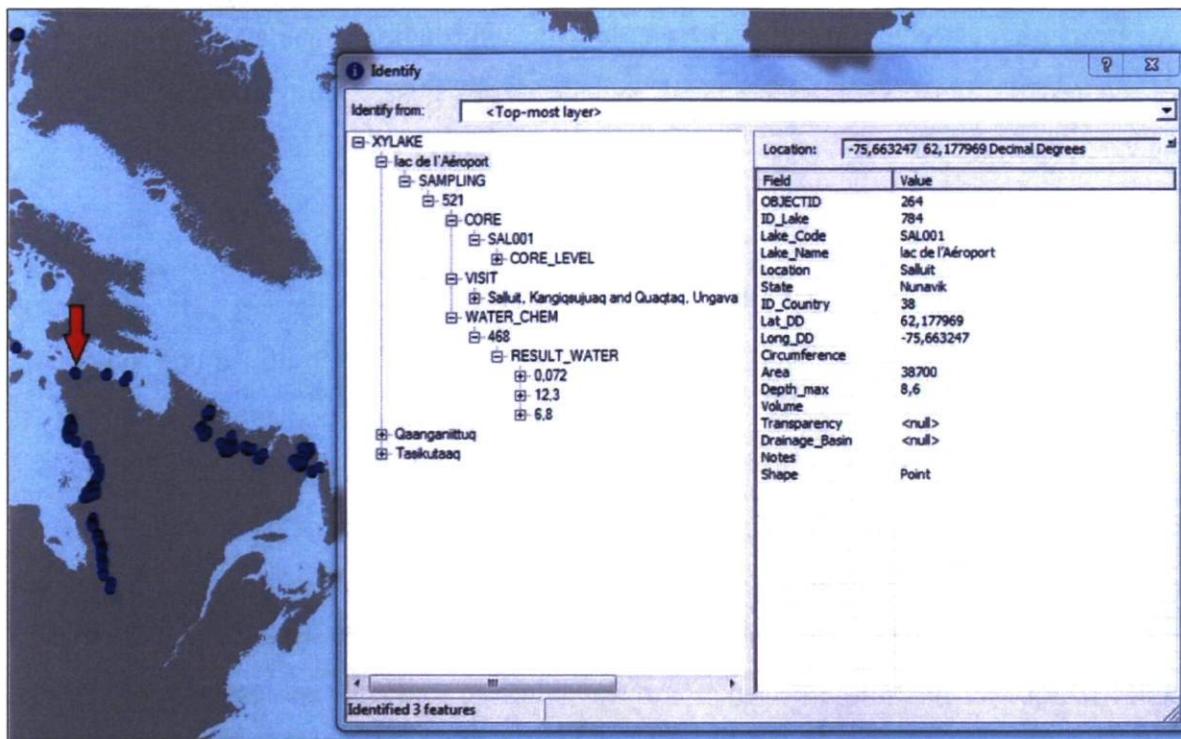


Figure 28 : Détail de l'utilisation de la CDD convertie en Personal Geodatabase dans ArcGIS.

L'extraction par requête de sélection

Tel que mentionné un peu plus haut, par rapport au traitement des données de la CDD dans un logiciel externe, il est également possible d'effectuer une sélection par requête sur n'importe quelles données de la CDD et de l'extraire sous format texte ou tableau MS Excel, par exemple. Ce potentiel offert par la nouvelle version de la CDD demeure l'une des façons les plus simples d'obtenir les données à analyser pour l'utilisateur moyen. L'extraction par requête est une méthode efficace et couramment employée dans plusieurs systèmes d'information, notamment au ministère des Ressources naturelles et de la Faune (MRNF) du Québec. Nous avons testé cette méthode dans ArcGIS en réalisant des représentations cartographiques sommaires, deux d'entre-elles étant présentées aux Figures

29 et 30. Ces tests de requête, ainsi que de nombreux autres effectués à la phase de réalisation technique, ont permis de confirmer que la base de données a été conçue de façon adéquate et que toutes les données qu'elle contient sont accessibles par requête.

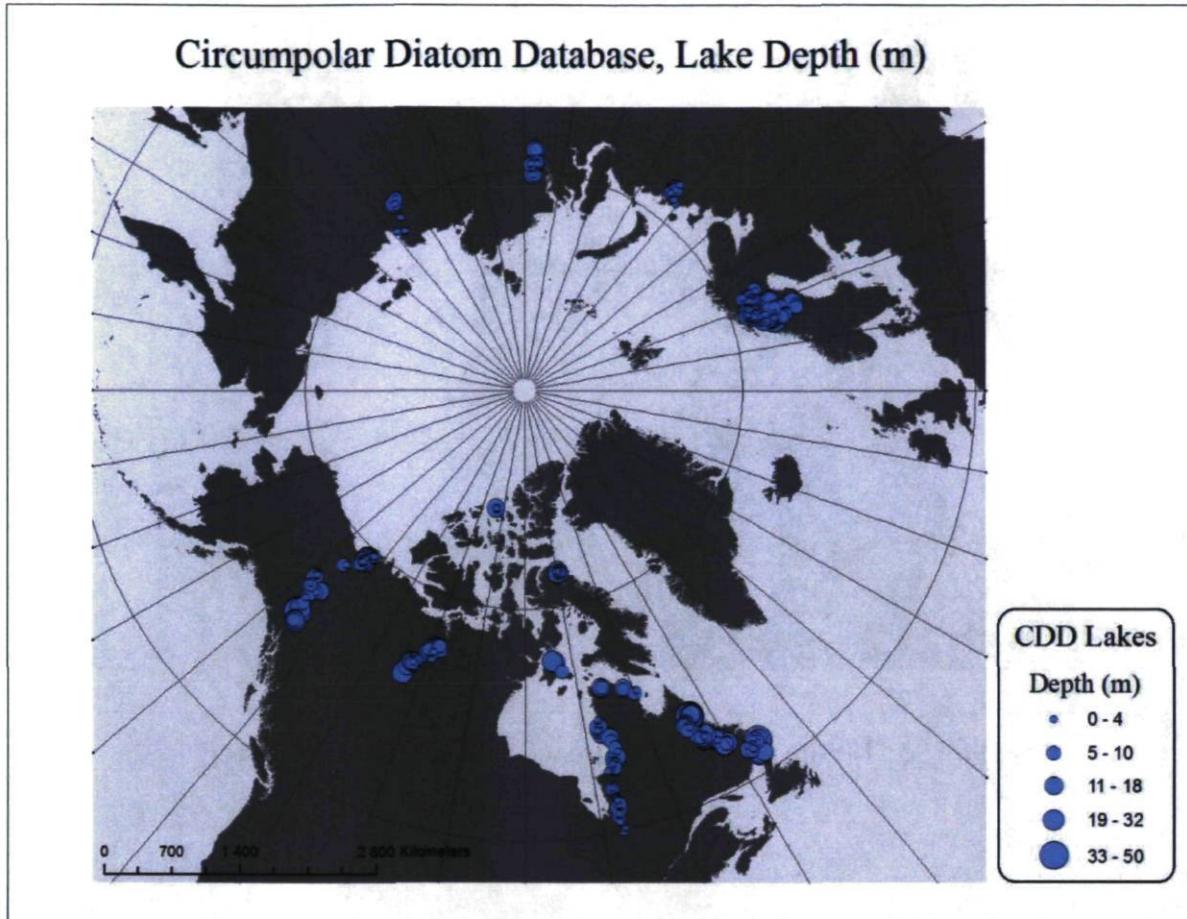


Figure 29 : Profondeurs des lacs de la CDD.

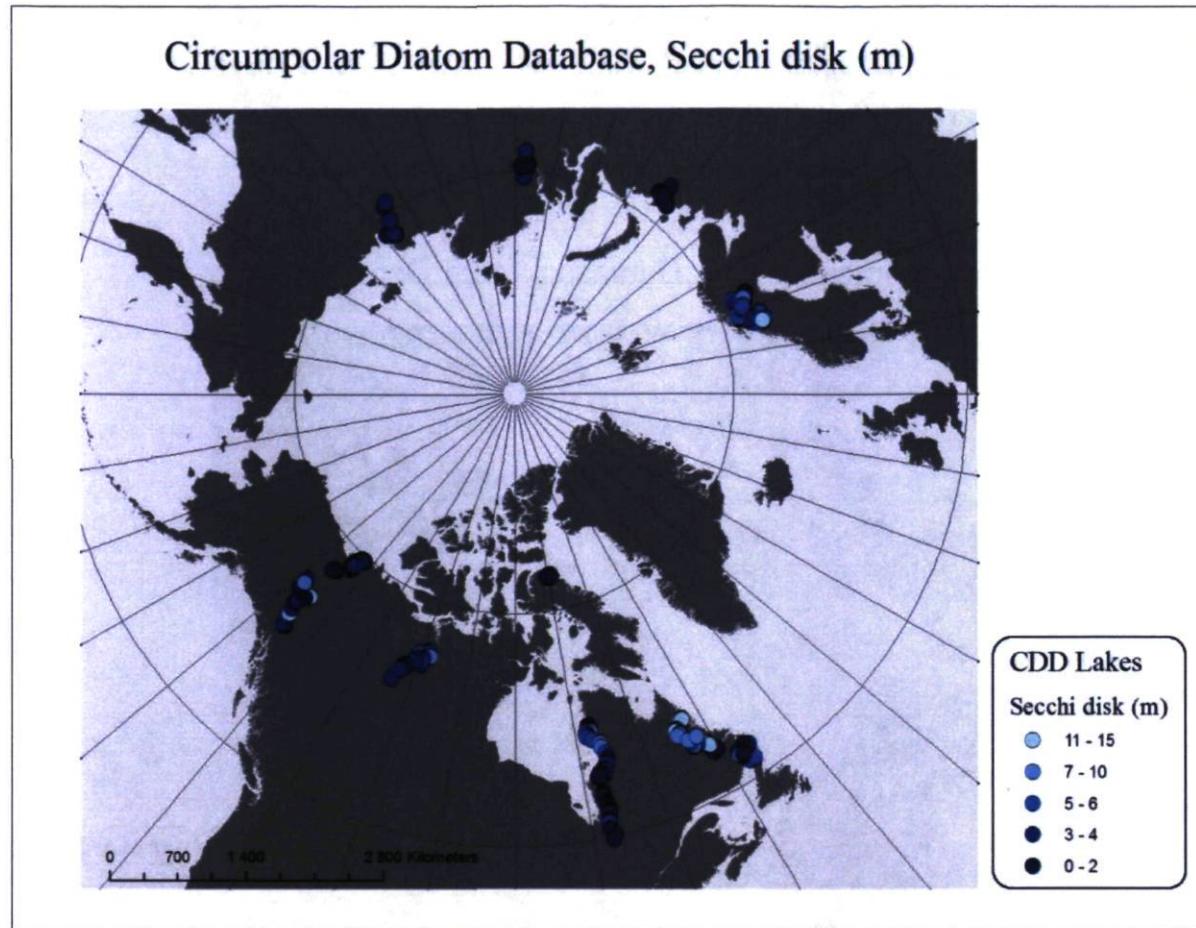


Figure 30 : Mesures de la transparence de l'eau par disque Secchi.

6. Chapitre 6 : Analyse et interprétation des résultats

À la lumière des résultats obtenus suite au développement de la base de données, des formulaires de consultation créés et des tests d'utilisation SIG effectués, il est possible de juger de la validité de la démarche entreprise et de faire ressortir les points forts, autant que les points faibles de cette nouvelle version de la Circumpolar Diatom Database. Les étapes restantes à développer ainsi que des suggestions d'améliorations à apporter au SI sont également soulevées, et ce, dans l'objectif d'encourager la continuité de cet ambitieux projet.

6.1. Fonctionnement de la structure du SI

Les requêtes effectuées par les tests de consultation et le développement des formulaires permettent d'affirmer que la structure de la base de données relationnelle est pleinement fonctionnelle. La structure des tables et de leurs relations est logique et permet d'accéder facilement à chacune des données du SI. Les attributs assignés aux champs ont été choisis conformément à la nature des données qu'ils contiennent. De cette manière, l'importation des données usuelles à la paléolimnologie peut être effectuée en minimisant les étapes de conversion préalables. La consultation des tables et des vues s'effectue en quelques fractions de seconde. Les requêtes qui prennent le plus de temps pour afficher ses résultats sont celles de graphiques et de tableaux croisés dynamiques, puisqu'elles traitent un grand nombre de données tout en les restructurant pour faciliter la consultation. Celles-ci s'affichent néanmoins dans des délais n'excédant pas 45 secondes. Les fichiers physiques étant divisés en trois parties (Données, Taxa et Formulaires) connectées entre elles par les options de liaison de tables, préviennent la corruption des données par des erreurs humaines ou logicielles. De plus, cette organisation des fichiers contribue à simplifier la distribution de la BD ainsi que sa gestion et les mises à jour.

La structure de la base de données présente toutefois une légère contrainte de confection. Il s'agit des deux tables de concordances « RESULT_PUB » et « PEOPLE_AFFIL » (Figure 10). Celles-ci étaient nécessaires pour gérer les objets qui étaient sollicités plusieurs fois dans différentes relations qui se recoupaient. Les tables de concordances sont fréquemment

employées dans les bases de données relationnelles et elles sont efficaces. Toutefois, leur utilisation n'est pas recommandée dans les systèmes simples puisqu'elles compliquent légèrement la création requêtes SQL ou les codes de programmation (pour modules ou macros). Néanmoins, cette difficulté, lorsqu'elle est connue, est facilement surmontable et concerne davantage le gestionnaire de base de données plutôt que l'utilisateur.

6.2. L'interface utilisateur

Les formulaires de consultation créés dans « CDD_Home.mdb » permettent de consulter aisément les résultats d'analyses et leurs informations associées. Les modules de recherche par taxon et par lac offrent aux utilisateurs différentes façons d'obtenir les résultats recherchés sans avoir à utiliser les requêtes SQL. La consultation par niveau de carottage est possible et les différents modes de présentation des données, par tableau et graphique croisés dynamiques, permettent de visionner les données d'une même carotte de façon cohérente. Cela répond directement à l'un des principaux objectifs de ce projet.

Quelques contraintes ont toutefois été constatées concernant l'affichage de résultats dans deux sections des formulaires. D'abord, dans le formulaire « Publication details » à la section du nom de la publication, la concaténation des champs nécessaires à l'obtention du nom complet n'est effective que lorsque la chaîne de caractères n'excède pas 255 caractères. De plus, les caractères spéciaux figurant parfois dans la nomenclature des numéros d'édition, les volumes ou les nombres de pages (ex. les deux-points ou les parenthèses) se confondent au langage de requête de concaténation et ne s'affichent pas correctement ou pas du tout. Pour résoudre ce problème, il pourrait être envisagé de créer un nouveau champ dans la table « PUB » du fichier « CDD.mdb ». Ce nouveau champ devrait avoir l'attribut « Mémo » (pouvant contenir plus de 255 caractères) de sorte qu'il pourrait contenir le nom complet de la publication. Cette option aurait pour avantage de contourner l'étape de concaténation, mais elle aurait comme inconvénient d'augmenter le volume de la table et de la BD, en plus d'accroître le risque d'erreurs de frappe. Une autre option serait d'élaborer un module de programmation dans « CDD_Home.mdb » afin d'éliminer les restrictions de la requête de concaténation. Cette option nécessiterait

toutefois de bonnes connaissances en programmation et la compatibilité du fichier CDD de formulaires pourrait être diminuée.

La mise en forme du graphique croisé dynamique d'abondances diatomifères présente aussi une complication dans le formulaire « Results by lake » à l'onglet « Diatom Graph ». En effet, il a été observé que l'orientation horizontale des diagrammes à bandes n'est pas conservée lors de l'ouverture des résultats pour un lac qui ne contient pas de valeur dans cette section. Il en résulte que le diagramme se présente dans sa position par défaut, c'est-à-dire en bandes verticales. Cette mise en page ainsi changée n'altère cependant pas les résultats, mais leur lecture devient plus ardue. Plusieurs essais ont été tentés pour résoudre ce problème de style, en passant par des changements dans les requêtes, les sous-formulaires et par le code VBA de MS Access, mais aucun résultat concluant n'a été obtenu. De plus amples recherches sur cette complication devront être réalisées. L'exploration plus poussée des possibilités qu'offrent les modules de programmation macro et VBA semble vraisemblablement l'avenue à emprunter.

Outre ces deux contraintes observées dans les formulaires MS Access, les formulaires constituent un moyen de consultation efficace et accessible à des utilisateurs moins expérimentés. D'un point de vue graphique, l'organisation des fenêtres et les styles choisis pour la mise en page sont sobres et de couleurs peu vives, et ce dans le but de prioriser l'attention sur les données et pour permettre l'impression en noir et blanc autant qu'en couleur. D'ailleurs, l'une des principales étapes restant à développer dans les formulaires de résultats est la création de boutons d'impression et d'extraction pour une sélection de résultats. Cette opération pourra s'effectuer à partir du mode création de MS Access et à l'aide du code VBA. Ces options ajoutées aux formulaires permettront aux utilisateurs d'enregistrer une série de résultats à partir de leur recherche par formulaires.

6.3. Bilan des différents tests d'utilisation

Des avantages et des inconvénients ont été constatés suite aux tests d'utilisation de la CDD dans ArcGIS. Ceux-ci sont résumés aux Tableaux 3 et 4. Bien que ces tests aient démontré qu'il existait certaines contraintes dans la consultation et le traitement des données de la

CDD par consultation directe (OLE DB, PersonalGDB), il demeure que le contenu de la CDD peut être extrait dans une multitude de formats permettant aux utilisateurs d'atteindre leurs objectifs d'analyse. Les conclusions tirées de ces essais et observations ont permis de cibler les causes possibles des problèmes de consultation directe avec ArcGIS. Ainsi, de plus amples lectures sur la compatibilité entre SGBD et les logiciels SIG et d'autres tests d'utilisation pourraient permettre une meilleure communication entre ces programmes. Entre autres, l'idée a été soulevée de créer un module de programmation permettant de traduire les ID entre la CDD et ArcGIS (Cournoyer; Girard; Bouchard; Janssens-Coron, communications personnelles, 2010). En d'autres mots, il s'agirait de placer la CDD sur un serveur informatique (par exemple avec Microsoft SQL Server qui est utilisé au CEN) et de programmer une composante dans ce serveur facilitant la connexion avec ArcGIS. Cette opération nécessiterait des compétences poussées en informatique lors de sa création, mais une fois celle-ci créée et le SI fonctionnel, sa gestion pourrait se faire par un gestionnaire moins qualifié. Une autre piste de solution qui peut être explorée est de recréer la base de données dans MS Access ou un autre SGBD, mais cette fois en obligeant la création d'ObjectID compatibles avec ArcGIS pour chaque table. Des tests supplémentaires devront être faits pour améliorer la connectivité directe avec les logiciels SIG et ceux-ci pourront guider les choix de développement futur du SI.

Tableau 3 : Avantages et inconvénients de la connexion OLE DB.

Avantages	Inconvénients
Simple	Hiérarchies des relations
Distribution par serveur	ID mal gérés
Sécuritaire	Sélection spatiale impossible
Relations fonctionnelles	Certains attributs incompatibles

Tableau 4 : Avantages et inconvénients de la conversion en PersonalGDB.

Avantages	Inconvénients
Hiérarchies des relations	Gestion plus complexe
Sélection spatiale	Relations non conservées
ID uniques	Moins accessible (ArcGIS)
	Compatibilité diminuée

Conclusion et recommandations

Les besoins sont nombreux en matière d'unification de données portant sur les diatomées circumpolaires et les conditions environnementales. L'amélioration de la Circumpolar Diatom Database pourrait aider à y répondre et c'est pourquoi le développement d'une nouvelle structure de base de données relationnelle a été exploré. Ce projet de recherche a permis de démontrer qu'à partir d'un système d'information spécialement adapté à la paléolimnologie, il est possible de consulter, d'archiver, d'analyser et de représenter efficacement les données de différents travaux de recherche. La démarche empruntée au cours de ce projet a nécessité une méthodologie étendue et les résultats engendrés ont été soumis à une analyse objective. De nombreuses recommandations ont été émises et des suggestions de perfectionnement ont été soulevées. Cette section offre une synthèse claire de ces recommandations afin d'aider la continuation éventuelle de ce projet. En effet, le SI développé offre de nombreuses possibilités d'exploitation et de diffusion qu'il est important de mettre en lumière. Préalablement à cette synthèse, un retour sur les résultats et leur interprétation a été dressé.

Les premières phases de développement d'un système d'information et l'évaluation des besoins du projet ont justifié l'emploi du modèle relationnel pour la création de la nouvelle base de données. Les phases de développement subséquentes ont mené à la création d'une structure de base de données finale permettant d'accéder à chacune des données par le biais de requêtes SQL et de formulaires de consultation. Les formulaires de consultation ont été développés pour que la navigation puisse s'effectuer par menus et fonctions automatisés, c'est-à-dire sans que l'utilisateur n'ait besoin d'utiliser le langage SQL. Cette interface permet ainsi d'accroître l'accessibilité de la CDD.

La structure physique de la CDD, en trois fichiers avec tables liées, préserve l'intégrité des données et favorise la diffusion du système d'information. L'élaboration de la méthode d'organisation de données par niveau, la méthode de standardisation des données sources, et la méthode de correspondance taxonomique ont rendu possible, respectivement, la consultation des données par niveau, l'importation par groupe de données normalisées et la création d'une librairie taxonomique pour la CDD. Conséquemment, la nouvelle version de

la CDD répond aux besoins d'archivage, d'analyse, de consultation et de représentation cartographique des données paléolimnologiques.

Grâce aux contributions de la communauté scientifique et certaines recherches récentes du LPA, la nouvelle version de la CDD a plus que triplé son volume d'informations. Elle renferme maintenant quelque 572 sites d'échantillonnage, 39 350 données d'occurrence et 15 000 données limnologiques. Sa couverture s'étend maintenant sur huit régions circumpolaires réparties sur trois continents : Amérique du Nord, Europe (Scandinavie) et Asie (Sibérie). Bien que le volume de la CDD ait considérablement augmenté suite à ce projet, plusieurs régions circumpolaires arctiques ne sont toujours pas suffisamment représentées. Jusqu'à maintenant, la couverture spatiale de la CDD est limitée, ce qui rend difficile l'étude de phénomènes biogéographiques de vaste échelle. Outre la localisation, les périodes d'analyse des données gagneraient également à être étendues, dans le but d'accroître la couverture temporelle de la CDD. Il est donc primordial d'encourager la communauté paléolimnologique à contribuer davantage au projet, et ce en fournissant autant des données récentes que des données anciennes. La publication d'articles ou de communiqués à propos du projet de la CDD pourrait accroître l'intérêt des chercheurs à y collaborer. Des réseaux sociaux et des listes d'envoi¹⁰ ont déjà employés pour la collecte de données et auraient avantage à être maintenus. De plus, l'ébauche d'une page Internet a déjà été créée¹¹ et son utilisation représenterait une bonne façon d'expliquer et d'illustrer les différentes fonctionnalités et possibilités de la CDD. Finalement, l'implantation et la diffusion du SI en réseau constitueraient un argument prépondérant afin d'inciter la communauté à participer au projet.

L'utilisation de la CDD dans ArcGIS a été testée par une connexion OLE DB et par conversion en « Personal Geodatabase ». Des avantages et des inconvénients ont été constatés autant pour l'une que pour l'autre de ces méthodes. De façon générale, l'utilisation de la base de données est possible dans un logiciel SIG aux fins d'analyse et de représentation cartographique. Cependant, des améliorations pourraient être apportées pour

¹⁰ Par exemple, The Paleolimnology Forum (<http://www.stpt.usf.edu/mariedin/paleolim.htm>) et DIATOM-L Research on the diatom algae (<https://listserv.indiana.edu/cgi-bin/wa-iub.exe?A0=DIATOM-L>).

¹¹ Laboratoire de paléoécologie aquatique : <http://www.cen.ulaval.ca/paleo/CDD/indexfr.html>.

simplifier sa connectivité. Entre autres, le test de connectivité OLE DB a dévoilé certaines lacunes de compatibilité des identifiants uniques MS Access, ArcGIS nécessitant impérativement des ObjectID pour permettre la sélection spatiale et pour conserver la hiérarchie des relations.

Certains problèmes et suggestions pour l'amélioration du SI ont été soulevés au cours de l'analyse des résultats. Parmi ceux-ci, on compte la concaténation des noms de publications, dans le formulaire « Publication », l'orientation du graphique croisé dynamique dans le formulaire « Results by lake », et la création de boutons avec fonction d'impression et d'extraction. À cela s'ajoute un autre aspect moins discuté, celui des données de datations. Les données accumulées dans la CDD ne contenaient pas suffisamment de résultats de datation pour que leur incorporation puisse être effectuée. Néanmoins, la collecte et l'intégration de ce type de données devront être préconisées dans les futurs développements et contributions. Les données de datation, ayant déjà des champs qui leur ont été alloués dans les tables de résultats par niveau, seront d'une grande utilité pour les interprétations spatio-temporelles.

À ce jour, la CDD pourrait être diffusée sous sa forme actuelle, ne nécessitant que quelques ajustements préalables pour résoudre les problèmes précédemment soulevés. En outre, un module d'automatisation pour les liaisons de tables devrait être développé et des dispositions légales devront être prises à l'égard des données issues de collaborateurs externes. Toutefois, grâce à la structure relationnelle qui a été développée, des options de diffusion par réseau s'offrent également à la CDD. La diffusion réseau nécessiterait plus de travail d'implantation, mais elle comporterait plusieurs avantages, dont l'accessibilité, la visibilité et la notoriété. Des réflexions et projections quant à l'avenir du SI devront être effectuées avant d'entreprendre concrètement l'étape d'implantation. Dépendant des ressources matérielles et humaines disponibles et des objectifs visés par le responsable de la CDD, la CDD pourrait être dotée d'un serveur réseau avec possibilité de diffusion sur Internet. Ce type de diffusion dans la communauté accroîtrait probablement la popularité de la CDD et, conséquemment, attirerait davantage de collaborateurs. Jusqu'à maintenant, la CDD est gérée à partir de Microsoft Access. Toutefois, la conversion de la base de données relationnelle vers d'autres SGBD est possible, notamment vers ArcGIS et Microsoft SQL

Server. Le développement futur du SI peut donc prendre plusieurs avenues. Ce projet de recherche fournit une foule d'outils pour la mise en application d'un SI fonctionnel qui, espérons-le, saura occuper une place importante dans l'avancement des connaissances sur les diatomées circumpolaires, les conditions environnementales et les écosystèmes lacustres nordiques. Souhaitons que, dans un avenir rapproché, les collaborations afflueront et que la Circumpolar Diatom Database devienne une référence principale auprès des membres de la communauté de paléolimnologues.

Bibliographie

Ouvrages et articles de références

- ACIA, 2005. Impacts of a Warming Arctic: Arctic Climate Impact Assessment. Cambridge University Press, New York, 1042 p.
- Antoniades, D., Douglas, M. S. V. & Smol, J. P., 2003a. Comparative physical and chemical limnology of two Canadian High Arctic regions: Alert (Ellesmere Island, NU) and Mould Bay (Prince Patrick Island, NWT). *Archiv für Hydrobiologie* 158: 485-516.
- Antoniades, D., Douglas, M. S. V. & Smol, J. P., 2003b. The physical and chemical limnology of 24 ponds and one lake from Isachsen, Ellef Ringnes Island, Canadian High Arctic. *International Review of Hydrobiology* 88: 519-38.
- Antoniades, D., 2004. The limnology, diatom autecology, and paleolimnology of lakes and ponds from Alert, Ellesmere Island, Isachsen, Ellef Ringnes Island, and Mould Bay, Prince Patrick Island, Canadian High Arctic. Ph.D. thesis, University of Toronto, Toronto, Canada, 569 p.
- Antoniades, D., Douglas, M.S.V. & Smol, J.P., 2004. Diatom species-environment relationships and inference models from Isachsen, Ellef Ringnes Island, Canadian High Arctic. *Hydrobiologia* 529: 1-18.
- Antoniades, D., Douglas, M.S.V. & Smol, J.P., 2005. Diatom autecology and inference model development from the Canadian High Arctic Archipelago. *Journal of Phycology* 41: 30-45.
- Battarbee, R. W., Juggins S., Gasse F., Anderson N. J., Bennion H., Cameron N. G., Ryves D. B., Pailles C., Chali F. & Telford N., 2001. European Diatom Database (EDDI): an Information System for Palaeoenvironmental Reconstruction. London, ECRC Research Report 81: 1-94.
- Bouchard, G. Gajewski, K. & Hamilton, P.B., 2004. Freshwater diatom biogeography in the Canadian Arctic Archipelago. *Journal of Biogeography* 31: 1955-1973.
- Bergeron, M., 1993. Vocabulaire de la géomatique. Cahiers de l'Office de la langue française, Les Publications du Québec, 41 p.
- Bigler, C. & Hall, R., 2002. Diatoms as indicators of climatic and limnological change in Swedish Lapland: a 100-lake calibration set and its validation for paleoecological reconstructions. *Journal of Paleolimnology* 27: 97-115.
- Bisson, B., 1997. Modèles de données : études conceptuelle et relationnelle. Paris, Economica, 112 p.

- Charles, D.F. & Whitehead, D.R., 1986. The PIRLA project: Paleoecological investigations of recent lake acidification. *Hydrobiologia* 143: 13-20.
- Codd, E.F., 1970. A Relational Model of Data for Large Shared Data Banks. *Communications of the Association for Computing Machinery (ACM)* vol. 13(6): 377-387.
- Codd, E.F., 1990. The relational model for database management - version 2. Addison-Wesley Publishing Company, 538 p. (ISBN 0-201-14192-2)
- Côté, G., Pienitz, R., Velle, G. & Wang, X., 2010. Impact of geese on the limnology of lakes and ponds from Bylot Island (Nunavut, Canada). *International Review of Hydrobiologia* 95(2): 105-129.
- Duff, K.E., Laing, T.E., Smol, J.P. & Lean, D.R.S., 1999. Limnological characteristics of lakes located across arctic treeline in northern Russia. *Hydrobiologia* 391: 205-222.
- Fallu, M.-A. & Pienitz, R., 1999. Diatomées lacustres de Jamésie-Hudsonie (Québec) et modèle de reconstitution des concentrations de carbone organique dissous. *Écoscience* 6(4): 603-620.
- Fallu, M.-A., Allaire, N. & Pienitz, R., 2000. Freshwater Diatoms from northern Québec and Labrador (Canada). *Bibliotheca Diatomologica*, vol. 45. J. Cramer, Berlin/Stuttgart, 200 p.
- Fallu, M.-A., Allaire, N. & Pienitz, R., 2002. Distribution of freshwater diatoms in 64 Labrador (Canada) lakes: species-environment relationships along latitudinal gradients and reconstruction models for water colour and alkalinity. *Canadian Journal of Fisheries and Aquatic Sciences* 59: 329-349.
- Gosselain, V., Coste, M., Campeau, S., Ector, L., Fauville, C., Delmas, F., Knoflacher, M., Licursi, M., Rimet, F., Tison, J., Tudesque, L. & Descy, J.P., 2005. A large-scale stream benthic diatom database. *Hydrobiologia* 542: 151-163.
- Hainaut, J.L., 2009. Bases de données : Concepts, utilisation et développement. Dunod, 704 p.
- Hartley, B., 1986. A check-list of the freshwater, brackish and marine diatoms of the British Isles and adjoining coastal waters. *Journal of the Marine Biological Association of the United Kingdom* 66: 531-610.
- Laing, T.E., Pienitz, R. & Smol, J.P., 1999a. Freshwater diatom assemblages from 23 lakes located near Norilsk, Siberia : A comparison with assemblages from other circumpolar treeline regions. *Diatom Research* 14(2): 285-305.

- Laing, T.E., Rühland, K.M. & Smol, J.P., 1999b. Past environmental and climatic changes related to tree-line shifts inferred from fossil diatoms from a lake near the Lena River Delta, Siberia. *The Holocene* 9(5): 547-557.
- Laing, T.E. & Smol, J.P., 2000. Factors influencing diatom distributions in circumpolar treeline lakes of northern Russia. *Journal of Phycology* 36: 1035-1048.
- Laperrière, L., 2007. Évolution postglaciaire du secteur sud-ouest du Bassin de Foxe, Île de Southampton, inférée par les assemblages fossiles de diatomées. Thèse de maîtrise, Département de Géographie, Université Laval, Québec, 86 p.
- Ministère des Communications, 1993. La géomatique et le développement d'un système d'information à référence spatiale, Ministère des communications, Gouvernement du Québec, Québec, Canada, 75 p.
- Munro, M.A.R., Kreiser, A.M., Battarbee, R.W., Juggins, S., Stevenson, A.C., Anderson, D.S., Anderson, N.J., Berge, F., Birks, H.J.B., Davis, R.B., Flower, R.J., Fritz, S.C., Haworth, E.Y., Jones, V.J., Kingston, J.C. & Renberg, I., 1990. Diatom quality control and data handling. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 327: 257-261.
- Parent, S., 1990. Dictionnaire des sciences de l'environnement. Broquet, Bibliothèque nationale du Québec, Ottawa, Canada, 748 p.
- Pienitz, R., Douglas, M.S.V. & Smol, J.P. (eds), 2004. Long-term environmental changes in Arctic and Antarctic lakes. Developments in paleoenvironmental research (DPER) Series, vol. 8, Springer-Verlag, Dordrecht/Berlin/Heidelberg/ New York, 562 p.
- Pienitz, R., Smol, J.P. & Birks, H.J.B., 1995. Assessment of freshwater diatoms as quantitative indicators of past climatic change in Yukon and Northwest Territories, Canada. *Journal of Paleolimnology* 13: 21-49.
- Pienitz, R., Smol, J.P. & Lean, D.R.S., 1997a. Physical and chemical limnology of 59 lakes located between the southern Yukon and Tuktoyaktuk Peninsula, Northwest Territories (Canada). *Canadian Journal of Fisheries and Aquatic Sciences* 54: 330-346.
- Pienitz, R., Smol, J.P. & Lean, D.R.S., 1997b. Physical and chemical limnology of 24 lakes located between Yellowknife and Contwoyto Lake, Northwest Territories (Canada). *Canadian Journal of Fisheries and Aquatic Sciences* 54: 347-358.
- Roche, S., 2009. SIG et analyse spatiale. Département des sciences géomatiques, Université Laval, Québec, Canada. Notes de cours.

- Rolland, N., Larocque, I., Francus, P., Pienitz, R. & Laperrière, L., 2008. Holocene climate inferred from biological (Diptera : Chironomidae) analyses in a Southampton island (Nunavut, Canada) lake. *The Holocene* 18(2): 229-241.
- Rolland, N., Larocque, I., Francus, P., Pienitz, R. & Laperrière, L., 2009. Evidence for a warmer period during the 12th and 13th centuries AD from chironomid assemblages in a Southampton Island, Nunavut, Canada. *Quaternary Research* 72: 27-37.
- Rosén, P., Hall, R., Korsman, T. & Renberg, I., 2000. Diatom transfer-functions for quantifying past air temperature, pH and total organic carbon concentration from lakes in northern Sweden. *Journal of Paleolimnology* 24: 109-123.
- Saulnier-Talbot, É., 2007. Impacts de l'évolution climatique postglaciaire sur les lacs de l'extrême nord de l'Ungava, Québec. Thèse de doctorat, Département de Géographie, Université Laval, Québec, 265 p.
- Saulnier-Talbot, É., Leng, M.J. & Pienitz, R., 2007. Recent climate and stable isotopes in modern surface waters of northernmost Ungava Peninsula, Canada. *Canadian Journal of Earth Sciences* 44: 171-180.
- Smol, J.P., Birks, H.J.B. & Last, W.M. (eds.), 2001. *Tracking Environmental Change Using Lake Sediments*. vol. 3: Terrestrial, Algal, and Siliceous Indicators. Kluwer Academic Publishers, Dordrecht, The Netherlands, 400 p.
- Sullivan, T.J. & Charles, D.F., 1994. The feasibility and utility of a paleolimnology/paleoclimate data cooperative for North America. *Journal of Paleolimnology* 10: 265-273.
- Thériault, M., 1996a. Systèmes d'information géographique. Concepts fondamentaux. 2^{ième} édition, Québec, LATIG, Département de géographie, Université Laval, 165 p.
- Thériault, M., 1996b. Systèmes d'information géographique : Notions de base. 2^{ième} édition, Québec, LATIG, Département de géographie, Université Laval, 367 p.
- Vincent, W.F. & Pienitz, R., 1996. Sensitivity of high-latitude freshwater ecosystems to global change: Temperature and solar ultraviolet radiation. *Geoscience Canada* 23: 231-236.
- Vyverman, W., Verleyen, E., Sabbe, K., Vanhoutte, K., Sterken, M., Hodgson, D.A., Mann, D.G., Juggins, S., Van De Vijver, B., Jones, V., Flower, R., Roberts, D., Chepurnov, V.A., Kilroy, C., Vanormelingen, P. & De Wever, A., 2007. Historical processes constrain patterns in global diatom diversity. *Ecology* 88(8): 1924-1931.
- Werner, D., 1977. Introduction with a note on taxonomy. In D. Werner (ed.), *The biology of diatoms (Botanical monographs 13)*. Blackwell Scientific Publicshers, Oxford 1-17.

Whitehead, D.R., Charles, D.F. & Golstein, R.A., 1990. The PIRLA project (Paleoecological Investigation of Recent Lake Acidification): an introduction to the synthesis of the project. *Journal of Paleolimnology* 3: 187-194.

Williams, D.M., Hartley, B., Ross, R., Munro, M.A.R., Juggins, S. & Battarbee, R.W., 1988. A coded checklist of British Diatoms. ENSIS Publishing, London. 74 p.

Sites Internet

ANSF, Academy of Natural Sciences in Philadelphia. NADED Master-list of names and taxa codes for diatoms and other algae: Names used in the North American Diatom Ecological Database. Phycology Section, Patrick Center for Environmental Research, The Academy of Natural Sciences, Philadelphia, PA. Retreived [2010] from [En ligne] <http://diatom.ansp.org/dpdc/taxalist.asp> (Page consultée le 07 avril 2008).

ESRI Support, 2010. ArcGIS 9.2 Desktop Help. [En ligne].
<http://webhelp.esri.com/arcgisdesktop/9.2/> (Page consultée le 15 août 2010).

Grimm, E., 2000. North American Pollen Database. IGBP PAGES/ World Data Centre for Paleoclimatology. NOAA/NGDC Paleoclimatology Program, Boulder, CO, USA.
[En ligne]. <http://www.ngdc.noaa.gov/paleo/pollen.html> (Page consultée le 29 novembre 2009).

HL7, 2010. Health Level Seven. [En ligne]. <http://www.hl7.de/> (Page consultée le 4 août 2010).

NRCAN-GSC. Geological Survey of Canada. Canadian Diatom Database, Ressources Naturelles Canada. [En ligne]. <http://gsc.nrcan.gc.ca/> (Page consultée le 30 avril 2009).

NRCAN-ATLAS. L'Atlas du Canada. Ressources Naturelles Canada. [En ligne].
<http://atlas.nrcan.gc.ca/> (Page consultée le 9 novembre 2009)

OQLF, Office québécois de la langue française, 2010. Le Grand Dictionnaire Terminologique. Gouvernement du Québec. Canada. [En ligne].
<http://www.oqlf.gouv.qc.ca/> (Page consultée le 15 octobre 2010)

Communications personnelles

Bouchard, M.A., 2010. Ministère des Ressource naturelles du Québec, [Communications personnelles]. Québec, Canada. Hiver 2010 à l'automne 2010.

- Cournoyer, L., 2010. Centre d'études nordiques, Université Laval, [Communications personnelles]. Québec, Canada. Hiver 2010 à l'automne 2010.
- Girard, É., 2010. Commission Géologique du Canada, [Communications personnelles]. Québec, Canada. Automne 2008 à l'automne 2010.
- Janssens-Coron, É., 2010. Département de géomatique, Université Laval, [Communications personnelles]. Québec, Canada. Automne 2008 à l'été 2010.

Annexe A
Liste détaillée des tables du fichier CDD.mdb

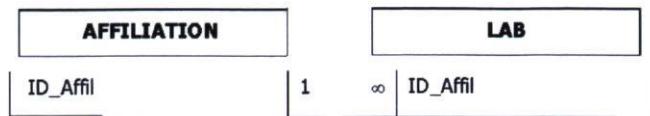
Propriétés

DateCreated:	2009-09-30 15:45:21	DefaultView:	2
DisplayViewsOnSharePointSite:	1	FilterOnLoad:	Faux
GUID:	{guid {6F05E1A8-F9B7-4124-8405-CCE026D0B75C}}	HideNewField:	Faux
LastUpdated:	2010-08-30 21:50:02	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	17
TotalsRow:	Faux	Updatable:	Vrai

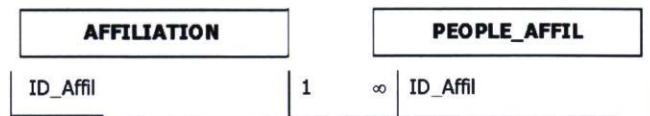
Colonnes

Nom	Type	Taille
ID_Affil	Entier long	4
Affil_Name	Texte	150
Code_Name	Texte	30
ID_Univ	Entier long	4
Full_Address	Texte	255
Street_Nbr	Entier	2
Street_Type	Texte	30
Street_Name	Texte	255
Pavillon	Texte	255
Floor	Entier	2
Room	Texte	30
Postal_code	Texte	20
Postal_case	Entier long	4
City	Texte	255
State	Texte	255
ID_Country	Entier long	4
Notes_1	Texte	255
Notes_2	Texte	255

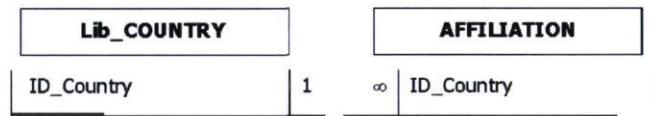
Relations

AFFILIATIONLAB

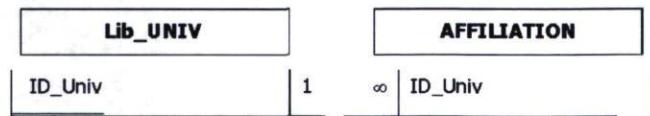
Attributes: Appliquée
RelationshipType: Un-à-plusieurs

AFFILIATIONPEOPLE_AFFIL

Attributes: Appliquée
RelationshipType: Un-à-plusieurs

Lib_COUNTRYAFFILIATION

Attributes: Appliquée
RelationshipType: Un-à-plusieurs

Lib_UNIVAFFILIATION

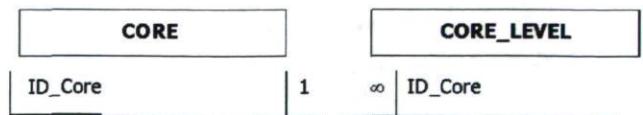
Attributes: Appliquée
RelationshipType: Un-à-plusieurs

Propriétés

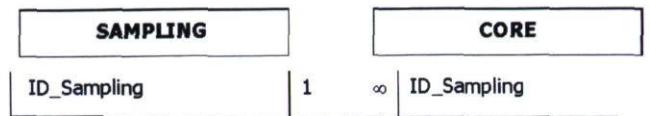
DateCreated:	2009-09-30 15:37:10	DefaultView:	2
DisplayViewsOnSharePointSite	1	FilterOnLoad:	Faux
GUID:	{guid {6168F5CD-0962-466E-9622-55B476753C6C}}	HideNewField:	Faux
LastUpdated:	2010-08-30 21:50:02	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	549
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

Nom	Type	Taille
ID_Core	Entier long	4
Core_Code	Texte	50
ID_Sampling	Entier long	4
Lat_DD	Réel double	8
Long_DD	Réel double	8
X_UTM	Réel double	8
Y_UTM	Réel double	8
Core_Length	Réel double	8
Coring_Depth	Réel double	8
Age_max_estimated	Réel double	8
Age_precision	Réel double	8
Notes_1	Texte	255
Notes_2	Texte	255

Relations**CORECORE_LEVEL**

Attributes: Appliquée
 RelationshipType: Un-à-plusieurs

SAMPLINGCORE

Attributes: Appliquée

RelationshipType: Un-à-plusieurs

D:\CDD.mdb
Table: CORE_LEVEL

16 février 2011

Propriétés

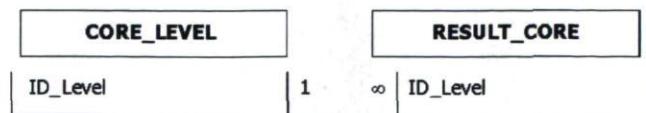
DateCreated:	2009-09-30 15:27:47	DefaultView:	2
DisplayViewsOnSharePointSite:	1	FilterOnLoad:	Faux
GUID:	{guid {45D5F46A-6170-4E3C-93DC-86F2463F30AA}}	HideNewField:	Faux
LastUpdated:	2010-08-30 21:50:02	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	1343
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

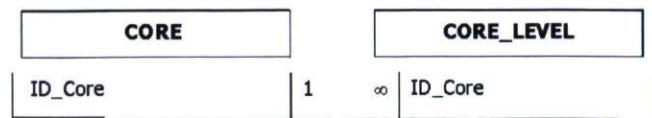
Nom	Type	Taille
ID_Level	Entier long	4
ID_Core	Entier long	4
Start_Depth	Réel double	8
End_Depth	Réel double	8
Weight_Total	Réel double	8
Weight_Chem	Réel double	8
Weight_Diatom	Réel double	8
Diatom_Total	Réel double	8
Microsphere_Total	Réel double	8
FOV	Réel double	8
Notes_1	Texte	255
Notes_2	Texte	255

Relations

CORE_LEVELRESULT_CORE



Attributes: Appliquée
RelationshipType: Un-à-plusieurs

CORECORE_LEVEL

Attributes: Appliquée
RelationshipType: Un-à-plusieurs

D:\CDD.mdb

16 février 2011

Table: INDEX_Taxa.Lib

Propriétés

DateCreated:	2010-05-30 13:29:58	DefaultView:	2
DisplayViewsOnSharePointSite:	1	Filter:	([INDEX_Taxa.Lib].[ID_Taxa.Lib]=207)
FilterOnLoad:	Faux	GUID:	{guid {3714868A-244E-451A-9B4D-4528F5A15A48}}
HideNewField:	Faux	LastUpdated:	2010-06-29 16:17:12
NameMap:	Donnée binaire	OrderBy:	[INDEX_Taxa.Lib].[ID_Diatom], [INDEX_Taxa.Lib].[ID_Taxa.Lib]
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	2472
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

Nom	Type	Taille
ID_Diatom	Entier long	4
ID_Taxa.Lib	Entier long	4
Taxa.Lib_Name	Texte	25

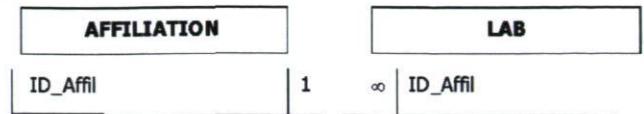
Propriétés

DateCreated:	2009-09-30 15:47:07	DefaultView:	2
DisplayViewsOnSharePointSite:	1	FilterOnLoad:	Faux
GUID:	{guid {37E5F9B9-250F-4B86-B61B-AB6EFA9A2C1E}}	HideNewField:	Faux
LastUpdated:	2010-08-30 21:50:02	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	6
TotalsRow:	Faux	Updatable:	Vrai

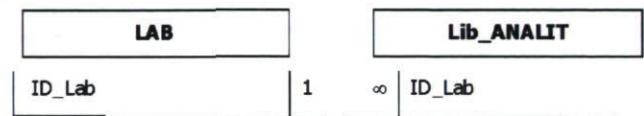
Colonnes

Nom	Type	Taille
ID_Lab	Entier long	4
Lab_Name	Texte	100
Code_Name	Texte	30
ID_Affil	Entier long	4
ID_Univ	Entier long	4
Full_Address	Texte	255
Street_Nbr	Entier	2
Street_Type	Texte	30
Street_Name	Texte	255
Pavillon	Texte	255
Floor	Entier	2
Room	Texte	30
Postal_code	Texte	30
Postal_case	Entier long	4
City	Texte	255
State	Texte	255
ID_Country	Entier long	4
Notes_1	Texte	255
Notes_2	Texte	255

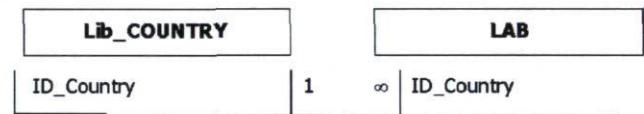
Relations

AFFILIATIONLAB

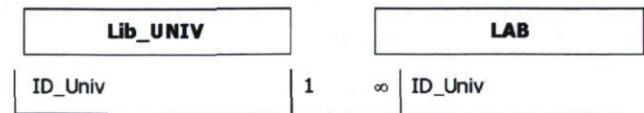
Attributes: Appliquée
 RelationshipType: Un-à-plusieurs

LIBLIB_ANALIT

Attributes: Appliquée
 RelationshipType: Un-à-plusieurs

LIB_COUNTRYLAB

Attributes: Appliquée
 RelationshipType: Un-à-plusieurs

LIB_UNIVLAB

Attributes: Appliquée
 RelationshipType: Un-à-plusieurs

Propriétés

DateCreated:	2008-11-25 14:39:08	DefaultView:	2
DisplayViewsOnSharePointSite:	1	FilterOnLoad:	Faux
GUID:	{guid {518A383D-355B-46A8-90F8-F8E8CCEB2E90}}	HideNewField:	Faux
LastUpdated:	2010-09-12 23:31:27	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	572
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

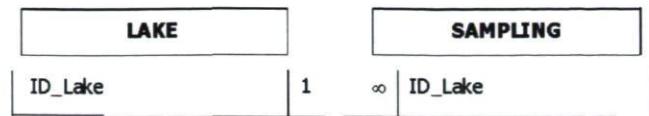
Nom	Type	Taille
ID_Lake	Entier long	4
Lake_Code	Texte	20
Lake_Name	Texte	255
Location	Texte	120
State	Texte	255
ID_Country	Entier long	4
Lat_DD	Réel double	8
Long_DD	Réel double	8
Elevation	Réel double	8
X_UTM	Réel double	8
Y_UTM	Réel double	8
Z_UTM	Réel double	8
Nbr_Core	Entier	2
Circumference	Réel double	8
Area	Réel double	8
Depth_avg	Réel double	8
Depth_max	Réel double	8
Volume	Réel double	8
Transparency	Réel double	8
Drainage_Basin	Réel double	8
Notes	Mémo	-

Relations

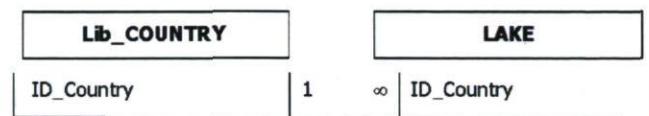
D:\CDD.mdb

16 février 2011

Table: LAKE

LAKESAMPLING

Attributes: Appliquée
RelationshipType: Un-à-plusieurs

Lib_COUNTRYLAKE

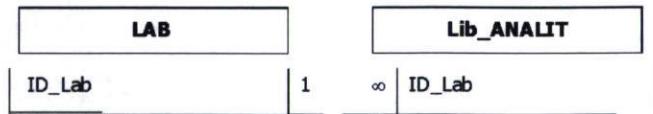
Attributes: Appliquée
RelationshipType: Un-à-plusieurs

Propriétés

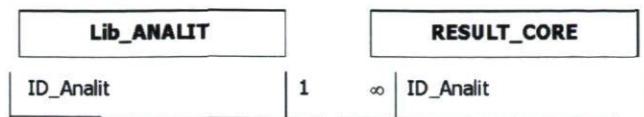
DateCreated:	2009-09-30 15:31:50	DefaultView:	2
DisplayViewsOnSharePointSite:	1	FilterOnLoad:	Faux
GUID:	{guid {6C628705-9CF9-434E-B4B0-9935E939A30B}}	HideNewField:	Faux
LastUpdated:	2010-08-30 21:50:02	NameMap:	Donnée binaire
OrderBy:	[Lib_ANALIT].[Name_Analit], [Lib_ANALIT].[ID_Analit]	OrderByOn:	Vrai
OrderByOnLoad:	Vrai	Orientation:	De gauche à droite
RecordCount:	92	TotalsRow:	Faux
Updatable:	Vrai		

Colonnes

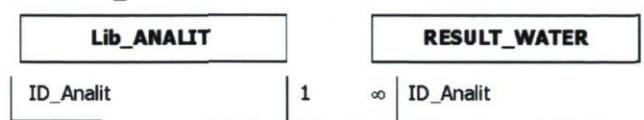
Nom	Type	Taille
ID_Analit	Entier long	4
Name_Analit	Texte	50
Definition	Mémo	-
ID_Units	Entier long	4
ID_Lab	Entier long	4
Category_Analit	Texte	50
Notes_1	Texte	255
Notes_2	Texte	255

Relations**LABLib_ANALIT**

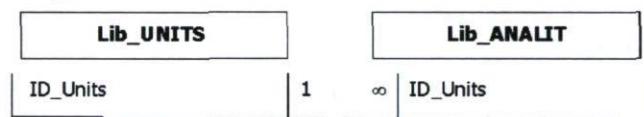
Attributes: Appliquée
RelationshipType: Un-à-plusieurs

Lib_ANALITRESULT_CORE

Attributes: Appliquée
RelationshipType: Un-à-plusieurs

Lib_ANALITRESULT_WATER

Attributes: Appliquée
RelationshipType: Un-à-plusieurs

Lib_UNITSlib_ANALIT

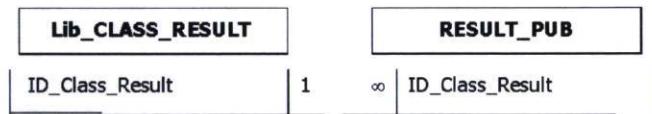
Attributes: Appliquée
RelationshipType: Un-à-plusieurs

Propriétés

DateCreated:	2009-10-23 15:00:07	DefaultView:	2
DisplayViewsOnSharePointSite:	1	FilterOnLoad:	Faux
GUID:	{guid {3F329894-E769-46F8-B66A-277E06CF70F6}}	HideNewField:	Faux
LastUpdated:	2010-08-30 21:50:02	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	0
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

Nom	Type	Taille
ID_Class_Result	Entier long	4
Class_Result	Texte	40

Relations**Lib_CLASS_RESULTRESULT_PUB**

Attributes: Appliquée
 RelationshipType: Un-à-plusieurs

D:\CDD.mdb
Table: Lib_COUNTRY

16 février 2011

Propriétés

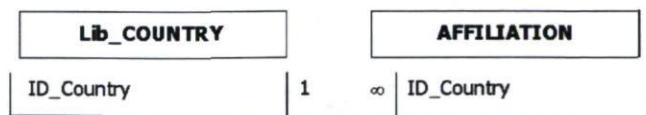
DateCreated:	2009-10-23 18:15:42	DefaultView:	2
DisplayViewsOnSharePointSite	1	FilterOnLoad:	Faux
GUID:	{guid {7C8D98AC-EA3A-4289-9826-ADA8B4EF0601}}	HideNewField:	Faux
LastUpdated:	2010-08-30 21:50:02	NameMap:	Donnée binaire
OrderBy:	[Lib_COUNTRY].[Country_name]	OrderByOn:	Vrai
OrderByOnLoad:	Vrai	Orientation:	De gauche à droite
RecordCount:	246	TotalsRow:	Faux
Updatable:	Vrai		

Colonnes

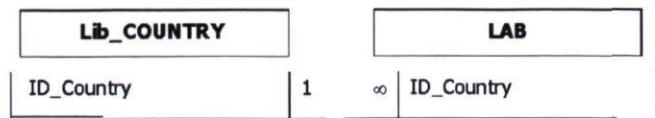
Nom	Type	Taille
ID_Country	Entier long	4
Country_names	Texte	100
ISO3166_1_alpha_2_code	Texte	2
French_names	Texte	100
ISO3166_alpha_3	Texte	3
ISO3166_number	Texte	3

Relations

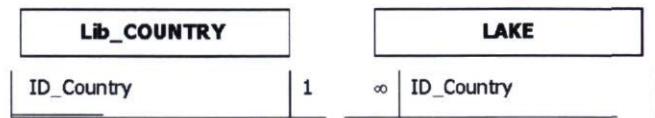
Lib_COUNTRYAFFILIATION



Attributes: Appliquée
RelationshipType: Un-à-plusieurs

Lib_COUNTRYLAB

Attributes: Appliquée
RelationshipType: Un-à-plusieurs

Lib_COUNTRYLAKE

Attributes: Appliquée
RelationshipType: Un-à-plusieurs

D:\CDD.mdb
Table: Lib_DIATOM

16 février 2011

Propriétés

DateCreated:	2008-12-05 14:47:24	DefaultView:	2
DisplayViewsOnSharePointSite:	1	Filter:	([Lib_DIATOM].[Genus] = "Misia")
FilterOnLoad:	Faux	GUID:	{guid {4C7BEB73-C26E-4798-B7-2507954E7E04}}
HideNewField:	Faux	LastUpdated:	2010-08-30 21:50:02
NameMap:	Donnée binaire	OrderBy:	[Lib_DIATOM].[ID_Diatom], [Lib_DIATOM].[Full_Taxon_Name], [Lib_DIATOM].[Genus], [Lib_DIATOM].[Species]
OrderByOn:	Vrai	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	7303
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

Nom	Type	Taille
ID_Diatom	Entier long	4
ID_CDD	Entier long	4
ID_Naded	Entier long	4
Genus	Texte	255
First_Taxon_Name	Texte	255
Species	Texte	255
Second_Taxon_Name	Texte	60
Third_Taxon_Name	Texte	255
Variety	Texte	255
Fourth_Taxon_Name	Texte	255
Form	Texte	255
Authority	Texte	255
Aut_Year	Entier	2
Short_Name	Texte	255
Full_Taxon_Name	Texte	255
FM_Code_Name	Texte	255
FM_Full_Tx_name	Texte	255
Notes_1	Texte	255
Notes_2	Texte	255

Relations

Lib_DIATOMRESULT_CORE

Lib_DIATOM	RESULT_CORE	
ID_Diatom	1	∞

Attributes: Appliquée
RelationshipType: Un-à-plusieurs

D:\CDD.mdb
Table: Lib_UNITS

16 février 2011

Propriétés

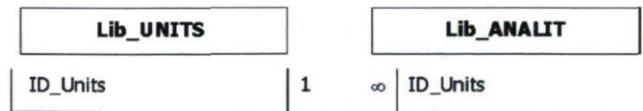
DateCreated:	2009-11-25 18:59:44	DefaultView:	2
DisplayViewsOnSharePointSite:	1	FilterOnLoad:	Faux
GUID:	{guid {F2E62290-E59B-450D-9046-28771CBFFA41}}	HideNewField:	Faux
LastUpdated:	2010-08-30 21:50:02	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	234
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

Nom	Type	Taille
ID_Units	Entier long	4
Unit_term	Texte	40
Alternatives	Texte	50
Name	Texte	255
Example_use	Texte	255
Canonical_form_value	Texte	150
Canonical_form_unit	Texte	120
cf_func	Texte	30

Relations

Lib_UNITSLib_ANALIT



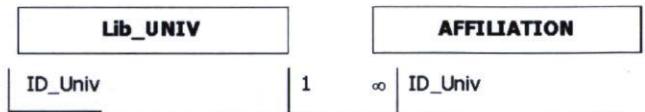
Attributes: Appliquée
RelationshipType: Un-à-plusieurs

Propriétés

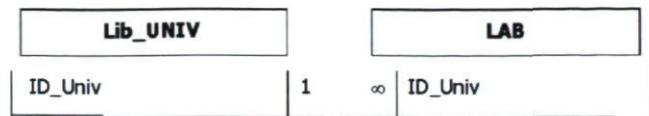
DateCreated:	2009-09-30 15:56:28	DefaultView:	2
DisplayViewsOnSharePointSite:	1	FilterOnLoad:	Faux
GUID:	{guid {45D0CD1F-1889-4EA1- ADC2-477FE4951420}}	HideNewField:	Faux
LastUpdated:	2010-08-30 21:50:02	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	3039
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

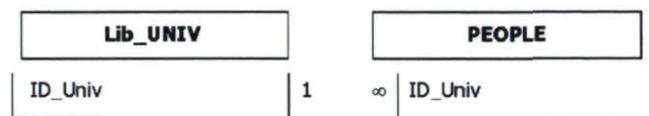
Nom	Type	Taille
ID_Univ	Entier long	4
Univ_Name	Texte	150
Full_Address	Texte	255
City	Texte	255
State	Texte	255
ID_Country	Entier long	4
Notes_1	Texte	255
Notes_2	Texte	255

Relations**Lib_UNIVAFFILIATION**

Attributes: Appliquée
 RelationshipType: Un-à-plusieurs

Lib_UNIVLAB

Attributes: Appliquée
RelationshipType: Un-à-plusieurs

Lib_UNIVPEOPLE

Attributes: Appliquée
RelationshipType: Un-à-plusieurs

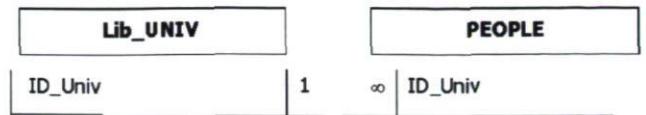
Propriétés

DateCreated:	2009-09-30 15:40:18	DefaultView:	2
DisplayViewsOnSharePointSite	1	FilterOnLoad:	Faux
GUID:	{guid {CE5AEB71-9950-4E49-A8D7-EDFE347DF759}}	HideNewField:	Faux
LastUpdated:	2010-08-30 21:50:02	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	13
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

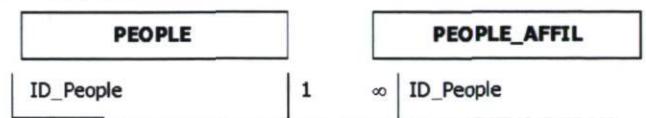
Nom	Type	Taille
ID_People	Entier long	4
Last_Name	Texte	100
First_Name	Texte	100
Title	Texte	100
ID_Univ	Entier long	4
Full_Address	Texte	150
Postal_Code	Texte	20
City	Texte	60
State	Texte	70
Country	Texte	70
Phone	Texte	60
Fax_Number	Texte	40
Email	Point d'ancre	-
Webpage	Point d'ancre	-
Notes_1	Texte	255
Notes_2	Texte	255

Relations

Lib_UNIVPEOPLE

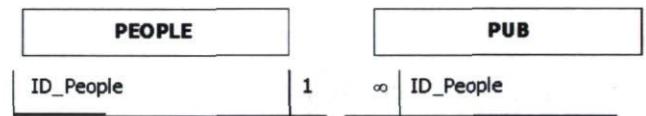
Attributes: Appliquée

RelationshipType: Un-à-plusieurs

PEOPLEPEOPLE_AFFIL

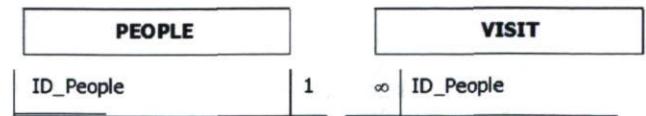
Attributes: Appliquée

RelationshipType: Un-à-plusieurs

PEOPLESUB

Attributes: Appliquée

RelationshipType: Un-à-plusieurs

PEOPLEVISIT

Attributes: Appliquée

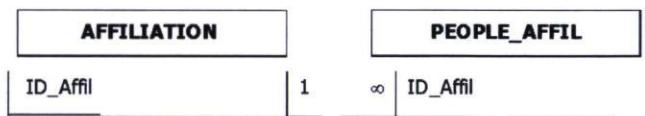
RelationshipType: Un-à-plusieurs

Propriétés

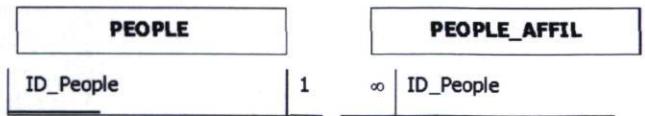
DateCreated:	2009-10-23 16:37:09	DefaultView:	2
DisplayViewsOnSharePointSite:	1	FilterOnLoad:	Faux
GUID:	{guid {CA2974A8-50FF-4DF5-B3FB-6C40E2D20E61}}	HideNewField:	Faux
LastUpdated:	2010-08-30 21:50:02	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	21
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

Nom	Type	Taille
ID_People	Entier long	4
ID_Affil	Entier long	4

Relations**AFFILIATIONPEOPLE_AFFIL**

Attributes: Appliquée
RelationshipType: Un-à-plusieurs

PEOPLEPEOPLE_AFFIL

Attributes: Appliquée
RelationshipType: Un-à-plusieurs

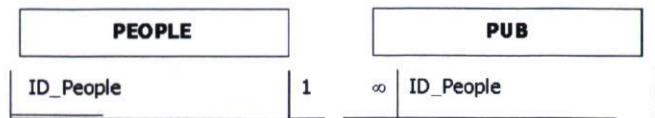
Propriétés

DateCreated:	2009-09-30 15:22:08	DefaultView:	2
DisplayViewsOnSharePointSite	1	FilterOnLoad:	Faux
GUID:	{guid {7340B219-0CE6-44B9-AE3C-6073DCF9AC34}}	HideNewField:	Faux
LastUpdated:	2010-08-30 21:50:02	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	21
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

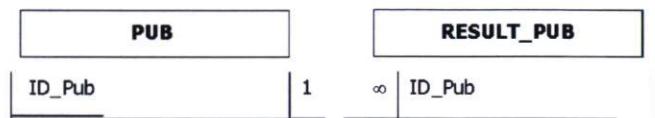
Nom	Type	Taille
ID_Pub	Entier long	4
Year_Pub	Entier	2
Nbr_Same_Year	Texte	2
First_Autor_Name	Texte	100
First_Autor_FirstName	Texte	100
ID_People	Entier long	4
Group_Autors	Mémo	-
Title	Mémo	-
Type	Texte	50
Pub_Name	Texte	255
Book_Chapter	Texte	10
Edition_site	Texte	255
Editor	Texte	255
Edition	Texte	15
Publisher	Texte	255
Volume	Texte	10
Number	Texte	5
Start_Page	Texte	15
End_Page	Texte	15
Total_Page	Entier	2
DOI_Name	Texte	255
ISBN	Texte	255
ISSN	Texte	255
Notes_1	Mémo	-
Notes_2	Texte	255

Relations

PEOPLEPUB

Attributes: Appliquée

RelationshipType: Un-à-plusieurs

PUBRESULT_PUB

Attributes: Appliquée

RelationshipType: Un-à-plusieurs

D:\CDD.mdb
Table: RESULT_CORE

16 février 2011

Propriétés

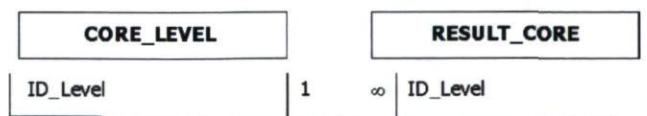
DateCreated:	2009-09-30 15:42:41	DefaultView:	2
DisplayViewsOnSharePointSite	1	FilterOnLoad:	Faux
GUID:	{guid {0A438DF3-7111-4819-BDA0-96A0678BB70F}}	HideNewField:	Faux
LastUpdated:	2010-08-30 21:50:02	NameMap:	Donnée binaire
OrderBy:	[RESULT_CORE].[Result]	OrderByOn:	Vrai
OrderByOnLoad:	Vrai	Orientation:	De gauche à droite
RecordCount:	41137	TotalsRow:	Faux
Updatable:	Vrai		

Colonnes

Nom	Type	Taille
ID_Result_Core	Entier long	4
ID_Level	Entier long	4
ID_Diatoms	Entier long	4
ID_Analit	Entier long	4
Result	Réel double	8
Notes_1	Texte	255
Notes_2	Texte	255

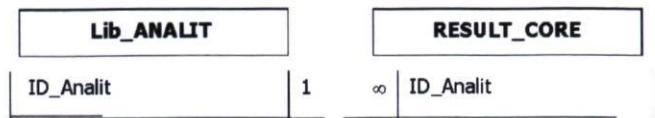
Relations

CORE_LEVELRESULT_CORE



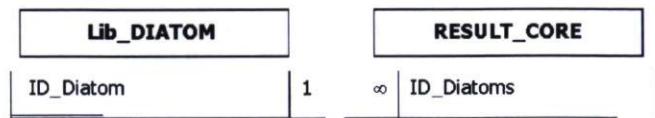
Attributes: Appliquée
RelationshipType: Un-à-plusieurs

Table: RESULT_CORE

Lib_ANALITRESULT_CORE

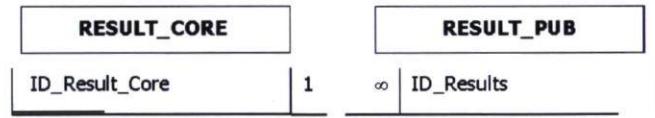
Attributes: Appliquée

RelationshipType: Un-à-plusieurs

Lib_DIATOMRESULT_CORE

Attributes: Appliquée

RelationshipType: Un-à-plusieurs

RESULT_CORERESULT_PUB

Attributes: Appliquée

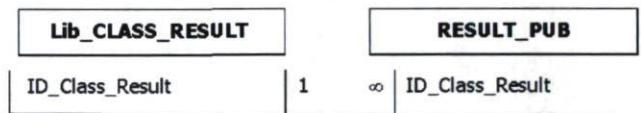
RelationshipType: Un-à-plusieurs

Propriétés

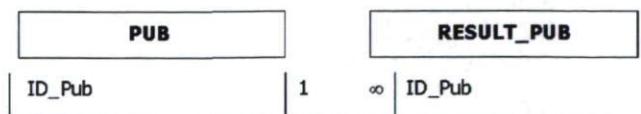
DateCreated:	2009-10-23 14:43:50	DefaultView:	2
DisplayViewsOnSharePointSite	1	FilterOnLoad:	Faux
GUID:	{guid {D78E74F5-D2AD-4233-903F-CBB4991F91B5}}	HideNewField:	Faux
LastUpdated:	2010-08-30 21:50:02	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	0
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

Nom	Type	Taille
ID_Class_Result	Entier long	4
ID_Results	Entier long	4
ID_Pub	Entier long	4

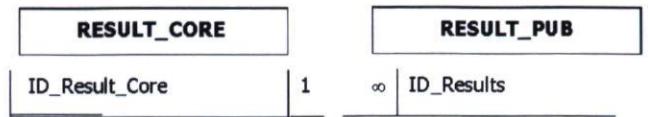
Relations**Lib_CLASS_RESULTRESULT_PUB**

Attributes: Appliquée
RelationshipType: Un-à-plusieurs

PUBRESULT_PUB

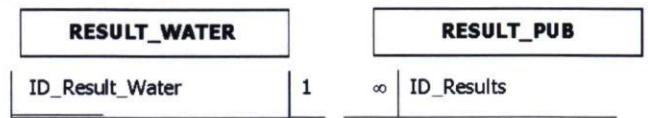
Attributes: Appliquée
RelationshipType: Un-à-plusieurs

Table: RESULT_PUB

RESULT_CORERESULT_PUB

Attributes: Appliquée

RelationshipType: Un-à-plusieurs

RESULT_WATERRESULT_PUB

Attributes: Appliquée

RelationshipType: Un-à-plusieurs

D:\CDD.mdb

16 février 2011

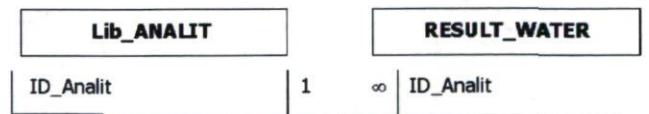
Table: RESULT_WATER

Propriétés

DateCreated:	2009-09-30 15:18:15	DefaultView:	2
DisplayViewsOnSharePointSite:	1	Filter:	((RESULT_WATER.ID_ResultWater>=549)) And ((RESULT_WATER.ID_ResultWater>=9282))
FilterOnLoad:	Faux	GUID:	{guid {3B5CB53E-AB97-43CA300-16EB16B69E74}}
HideNewField:	Faux	LastUpdated:	2010-09-10 21:30:23
NameMap:	Donnée binaire	OrderByOn:	Faux
OrderByOnLoad:	Vrai	Orientation:	De gauche à droite
RecordCount:	15274	TotalsRow:	Faux
Updatable:	Vrai		

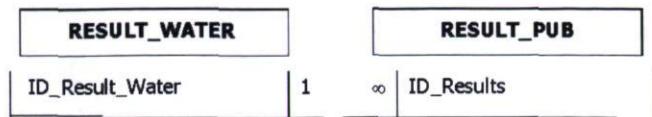
Colonnes

Nom	Type	Taille
ID_Result_Water	Entier long	4
ID_Water_Chem	Entier long	4
ID_Analit	Entier long	4
Result	Texte	100
Notes_1	Texte	255
Notes_2	Texte	255

Relations**Lib_ANALITRESULT_WATER**

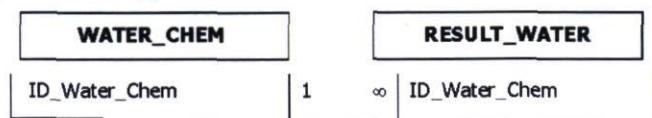
Attributes: Appliquée
 RelationshipType: Un-à-plusieurs

Table: RESULT_WATER

RESULT_WATERRESULT_PUB

Attributes: Appliquée

RelationshipType: Un-à-plusieurs

WATER_CHEMRESULT_WATER

Attributes: Appliquée

RelationshipType: Un-à-plusieurs

D:\CDD.mdb
Table: SAMPLING

16 février 2011

Propriétés

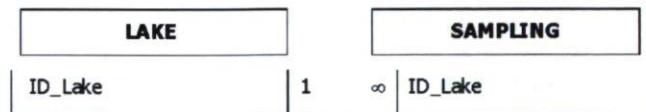
DateCreated:	2009-10-13 15:56:16	DefaultView:	2
DisplayViewsOnSharePointSite	1	FilterOnLoad:	Faux
GUID:	{guid {4A09216A-AAED-40CF-8821-03FFD0867257}}	HideNewField:	Faux
LastUpdated:	2010-08-30 21:50:02	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	833
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

Nom	Type	Taille
ID_Sampling	Entier long	4
Date_Sampling	Date/Heure	8
ID_Lake	Entier long	4
ID_Visit	Entier long	4
Notes_1	Texte	255
Notes_2	Texte	255

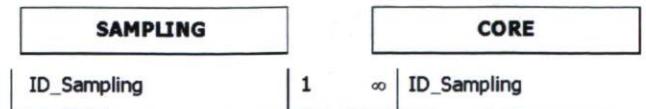
Relations

LAKESAMPLING

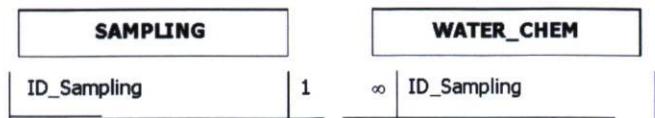


Attributes: Appliquée
RelationshipType: Un-à-plusieurs

SAMPLINGCORE

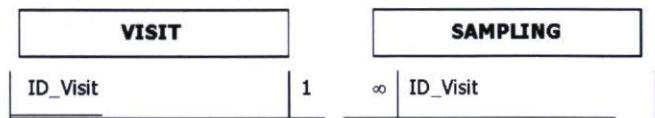


Attributes: Appliquée
RelationshipType: Un-à-plusieurs

SAMPLINGWATER_CHEM

Attributes: Appliquée

RelationshipType: Un-à-plusieurs

VISITSAMPLING

Attributes: Appliquée

RelationshipType: Un-à-plusieurs

D:\CDD.mdb

16 février 2011

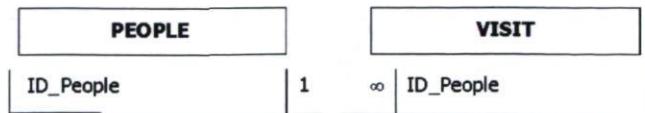
Table: VISIT

Propriétés

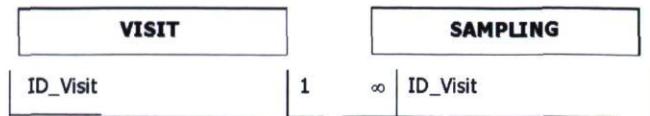
DateCreated:	2009-09-30 14:00:41	DefaultView:	2
DisplayViewsOnSharePointSite	1	FilterOnLoad:	Faux
GUID:	{guid {FDAB8045-9CCA-4148-9897-3EEF99D1C7A8}}	HideNewField:	Faux
LastUpdated:	2010-08-30 21:50:02	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	15
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

Nom	Type	Taille
ID_Visit	Entier long	4
Related_code	Texte	20
Start_Date	Date/Heure	8
End_Date	Date/Heure	8
ID_People	Entier long	4
Budget	Texte	50
Main_Rsrh_Prog	Mémo	-
Max_Lat_DD	Réel double	8
Min_Lat_DD	Réel double	8
West_Long_DD	Réel double	8
East_Long_DD	Réel double	8
Region_name	Texte	255
Notes_1	Texte	255
Notes_2	Texte	255

Relations**PEOPLEVISIT**

Attributes: Appliquée
 RelationshipType: Un-à-plusieurs

VISITSAMPLING

Attributes: Appliquée
RelationshipType: Un-à-plusieurs

D:\CDD.mdb
Table: WATER_CHEM

16 février 2011

Propriétés

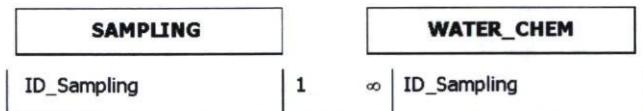
DateCreated:	2009-09-30 15:16:20	DefaultView:	2
DisplayViewsOnSharePointSite:	1	FilterOnLoad:	Faux
GUID:	{guid {CFD3197A-1096-4496-859D-31828762DF2F}}	HideNewField:	Faux
LastUpdated:	2010-08-30 21:50:02	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	568
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

Nom	Type	Taille
ID_Water_Chem	Entier long	4
ID_Sampling	Entier long	4
Notes	Mémo	-

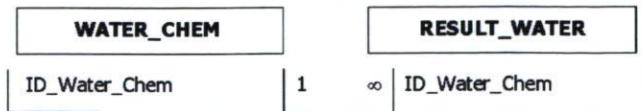
Relations

SAMPLINGWATER_CHEM



Attributes: Appliquée
RelationshipType: Un-à-plusieurs

WATER_CHEMRESULT_WATER



Attributes: Appliquée
RelationshipType: Un-à-plusieurs

Annexe B
Liste détaillée des tables du fichier CDD_Taxa_Lib.mdb

D:\CDD_Taxa.Lib.mdb

16 février 2011

Table: Abisko

Propriétés

DateCreated:	2010-05-30 13:08:47	DefaultView:	2
DisplayViewsOnSharePointSite:	1	FilterOnLoad:	Faux
GUID:	{guid {BC2C4CED-3D03-4181-9AE2-BD336348868F}}	HideNewField:	Faux
LastUpdated:	2010-05-30 13:08:47	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	293
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

Nom	Type	Taille
ID_Abisko	Réel double	8
Abisko_full_tx_name	Texte	255
Abisko_Genus	Texte	255
Abisko_FirstTxName	Texte	255
Abisko_Species	Texte	255
Abisko_SecondTxName	Texte	255
Abisko_ThirdTxName	Texte	255
Abisko_Variety	Texte	255
Abisko_FourthTxName	Texte	255
Abisko_Form	Texte	255
Abisko_UMEANORDCHILL_name	Texte	255
Syn_Krammer_LangeBertalot	Texte	255
Abisko_code	Texte	255
Abisko_Authority	Texte	255
Abisko_Aut_Year	Réel double	8

Propriétés

DateCreated:	2010-07-19 16:10:04	DefaultView:	2
DisplayViewsOnSharePointSite:	1	FilterOnLoad:	Faux
GUID:	{guid {DA6F8C2F-B608-470A-8503-AD0FC12D0D23}}	HideNewField:	Faux
LastUpdated:	2010-07-19 16:11:49	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	85
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

Nom	Type	Taille
ID	Entier long	4
BI_FullTxName	Texte	255
BI_TaxaCode	Texte	255
Genus	Texte	255
First_Taxon_Name	Texte	255
Species	Texte	255
Second_Taxon_Name	Texte	255
Third_Taxon_Name	Texte	255
Variety	Texte	255

Propriétés

DateCreated:	2010-05-30 13:09:27	DefaultView:	2
DisplayViewsOnSharePointSite:	1	FilterOnLoad:	Faux
GUID:	{guid {FE559F26-99CB-4005-895E-D6744B26DFDC}}	HideNewField:	Faux
LastUpdated:	2010-06-28 12:49:25	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	611
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

Nom	Type	Taille
ID_FM	Réel double	8
FM_full_TX_Name	Texte	255
FM_Genus	Texte	255
FM_FirstTxName	Texte	255
FM_Species	Texte	255
FM_SecondTxName	Texte	255
FM_ThirdTxName	Texte	255
FM_Variety	Texte	255
FM_FourthTxName	Texte	255
FM_Form	Texte	255
Taxon_code	Texte	255
Taxon_Name	Texte	255
FM_Authority	Texte	255
FM_Aut_Year	Réel double	8

Propriétés

DateCreated:	2010-05-30 13:08:24	DefaultView:	2
DisplayViewsOnSharePointSite:	1	FilterOnLoad:	Faux
GUID:	{guid {6D6971BF-5255-4398-9A6C-F02866BAECE6}}	HideNewField:	Faux
LastUpdated:	2010-05-30 13:08:24	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	246
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

Nom	Type	Taille
ID_Dermot	Réel double	8
Dermot_full_tx_name	Texte	255
DERMOT_Genus	Texte	255
DERMOT_FirstTxName	Texte	255
DERMOT_Species	Texte	255
DERMOT_SecondTxName	Texte	255
DERMOT_ThirdTxName	Texte	255
DERMOT_Variety	Texte	255
DERMOT_FourthTxName	Texte	255
DERMOT_Form	Texte	255
Dermot_Code	Texte	255

Propriétés

DateCreated:	2010-05-30 13:08:04	DefaultView:	2
DisplayViewsOnSharePointSite	1	FilterOnLoad:	Faux
GUID:	{guid {6A11BEA8-F713-4CE0-97FB-698574186648}}	HideNewField:	Faux
LastUpdated:	2010-05-30 13:08:04	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	268
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

Nom	Type	Taille
ID_Emilie	Réel double	8
Emilie_Full_TX_Name	Texte	255
EMILIE_Genus	Texte	255
EMILIE_FirstTxName	Texte	255
EMILIE_Species	Texte	255
EMILIE_SecondTxName	Texte	255
EMILIE_ThirdTxName	Texte	255
EMILIE_Variety	Texte	255
EMILIE_FourthTxName	Texte	255
EMILIE_Form	Texte	255
Emilie_Code	Texte	255

Propriétés

DateCreated:	2010-05-30 13:09:07	DefaultView:	2
DisplayViewsOnSharePointSite:	1	FilterOnLoad:	Faux
GUID:	{guid {07624C84-2B74-44AE-81AD-C86088DB9A21}}	HideNewField:	Faux
LastUpdated:	2010-05-30 13:16:03	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	739
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

Nom	Type	Taille
Diat code (UmU)	Réel double	8
DIAT_Sw_full_tx_name	Texte	255
DIAT_UME_Genus	Texte	255
DIAT_UME_FirstTxName	Texte	255
DIAT_UME_Species	Texte	255
DIAT_UME_SecondTxName	Texte	255
DIAT_UME_ThirdTxName	Texte	255
DIAT_UME_Variety	Texte	255
DIAT_UME_FourthTxName	Texte	255
DIAT_UME_Form	Texte	255
Ume namn	Texte	255
Nord-chill namn	Texte	255
Auktor (Ume)	Texte	255
Synonym (Ume)	Texte	255
Bokst kod Ume	Texte	255
2 bokst kod Nord-chill	Texte	255
Bokst kod Nord-chill	Texte	255
Amalgamering Nord-chill	Texte	255
Auktor nord-chill	Texte	255
årtal	Réel double	8

Propriétés

DateCreated:	2010-05-30 13:06:47	DefaultView:	2
DisplayViewsOnSharePointSite	1	FilterOnLoad:	Faux
GUID:	{guid {BAAAB49D-7687-4123-B858-CA024ECB0E5B}}	HideNewField:	Faux
LastUpdated:	2010-05-30 13:06:47	NameMap:	Donnée binaire
OrderByOn:	Faux	OrderByOnLoad:	Vrai
Orientation:	De gauche à droite	RecordCount:	214
TotalsRow:	Faux	Updatable:	Vrai

Colonnes

Nom	Type	Taille
ID_YEL	Réel double	8
YEL_full_taxon_name	Texte	255
YEL_Genus	Texte	255
YEL_FirstTxName	Texte	255
YEL_Species	Texte	255
YEL_SecondTxName	Texte	255
YEL_ThirdTxName	Texte	255
YEL_Variety	Texte	255
YEL_FourthTxName	Texte	255
YEL_Form	Texte	255
YEL_Code	Texte	255

Annexe C
Liste détaillée des tables du fichier CDD_Home.mdb

Objets**Section: Détail****Section: EntêteFormulaire****Section: PiedFormulaire****Bouton de commande: cmdClose****Bouton de commande: cmdGoToWaterResult****Étiquette: Étiquette24****Sous-formulaire/Sous-état: Fille13****Sous-formulaire/Sous-état: Fille21****Étiquette: lblNoResults****Sous-formulaire/Sous-état: sfrmResultDiatoms****Code**

```
VERSION 1.0 CLASS
BEGIN
    MultiUse = -1 'True
END
Attribute VB_Name = "Form_F_Abund_by_Diatom"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = True
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Compare Database
Option Explicit

Private Sub cmdGoToWaterResult_Click()
    Dim IngLakeID As Long

    IngLakeID = Me.sfrmResultDiatoms!ID_Lake

    DoCmd.OpenForm "F_Result_by_Lake"
    Form_F_Result_by_Lake.Search (IngLakeID)
End Sub
```

Objets**Section:** Détail**Rectangle:** Boîte37**Bouton de commande:** cmdSearch**Étiquette:** Étiquette12**Étiquette:** Étiquette32**Étiquette:** Étiquette35**Étiquette:** Étiquette4**Étiquette:** Étiquette6**Zone de liste:** lstDiatoms**Zone de liste:** lstFourthTxName**Zone de liste:** lstGenus**Zone de liste:** lstSpecies**Zone de liste:** lstThirdTxName**Code**

```
VERSION 1.0 CLASS
BEGIN
    MultiUse = -1 'True
END
Attribute VB_Name = "Form_F_Abund_by_Diatom1"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = True
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Compare Database
Option Explicit
```

```
Private Sub cmdSearch_Click()
```

D:\CDD_Home.mdb

16 février 2011

Formulaire: F_Abund_by_Diatom1

```
Form_F_Abund_by_Diatom.sfrmResultDiatoms.Form.RecordSource = "SELECT * FROM
r_Abund_by_Diatom WHERE ID_Diatoms = " & Me.lstDiatoms.Column(1,
Me.lstDiatoms.ListIndex)
```

```
Dim rs As Recordset
Set rs = Form_F_Abund_by_Diatom.sfrmResultDiatoms.Form.RecordsetClone
Form_F_Abund_by_Diatom.cmdGoToWaterResult.Enabled = rs.RecordCount > 0
Form_F_Abund_by_Diatom.lblNoResults.Visible = rs.RecordCount = 0
Form_F_Abund_by_Diatom.sfrmResultDiatoms.Visible = rs.RecordCount > 0
Set rs = Nothing
```

End Sub

```
Private Sub IstDiatoms_Click()
    Me.cmdSearch.Enabled = True
End Sub
```

```
Private Sub IstGenus_AfterUpdate()

If Me.lstGenus.ListIndex <> -1 Then
    Dim strSqlTemp As String
    strSqlTemp = "SELECT DiatomsWithResults.Species " & _
        "FROM DiatomsWithResults " & _
        "WHERE DiatomsWithResults.Genus "
```

```
'----- Genus -----
strSqlTemp = strSqlTemp & SqlWhere_Genus()
'-----
```

```
strSqlTemp = strSqlTemp & " GROUP BY DiatomsWithResults.Species ORDER BY
DiatomsWithResults.Species; "
```

```
Me.lstSpecies.RowSource = strSqlTemp
Me.lstSpecies.Requery
End If
```

```
IstSpecies_AfterUpdate
```

```
End Sub
```

```
Private Sub IstSpecies_AfterUpdate()
```

```
If Me.lstSpecies.ListIndex <> -1 Then
```

```

Dim strSqlTemp As String
strSqlTemp = "SELECT DiatomsWithResults.Third_Taxon_Name, " & _
    "DiatomsWithResults.Variety " & _
    "FROM DiatomsWithResults " & _
    "WHERE DiatomsWithResults.Genus "


'----- Genus -----
strSqlTemp = strSqlTemp & SqlWhere_Genus()
'-----



'----- Species -----
strSqlTemp = strSqlTemp & " AND DiatomsWithResults.Species "
strSqlTemp = strSqlTemp & SqlWhere_Species()
'-----



strSqlTemp = strSqlTemp & " GROUP BY DiatomsWithResults.Third_Taxon_Name,
DiatomsWithResults.Variety ORDER BY DiatomsWithResults.Third_Taxon_Name,
DiatomsWithResults.Variety; "


Me.IstThirdTxName.RowSource = strSqlTemp
Me.IstThirdTxName.Requery
End If

IstThirdTxName_AfterUpdate
End Sub

Private Sub IstThirdTxName_AfterUpdate()

If Me.IstThirdTxName.ListIndex <> -1 Then

Dim strSqlTemp As String
strSqlTemp = "SELECT DiatomsWithResults.Fourth_Taxon_Name, " & _
    "DiatomsWithResults.Form " & _
    "FROM DiatomsWithResults " & _
    "WHERE DiatomsWithResults.Genus "


'----- Genus -----
strSqlTemp = strSqlTemp & SqlWhere_Genus()
'-----



'----- Species -----
strSqlTemp = strSqlTemp & " AND DiatomsWithResults.Species "
strSqlTemp = strSqlTemp & SqlWhere_Species()
'-----
```

```
'----- Third Taxon Name -----
strSqlTemp = strSqlTemp & " AND DiatomsWithResults.Third_Taxon_Name "
strSqlTemp = strSqlTemp & SqlWhere_ThirdTxName()

'-----

'----- Variety -----
strSqlTemp = strSqlTemp & " AND DiatomsWithResults.Variety "
strSqlTemp = strSqlTemp & SqlWhere_Variety()

'-----



strSqlTemp = strSqlTemp & " GROUP BY DiatomsWithResults.Fourth_Taxon_Name,
DiatomsWithResults.Form ORDER BY DiatomsWithResults.Fourth_Taxon_Name,
DiatomsWithResults.Form; "

Me.lstFourthTxName.RowSource = strSqlTemp
Me.lstFourthTxName.Requery
End If

lstFourthTxName_AfterUpdate
End Sub

Private Sub lstFourthTxName_AfterUpdate()
    Recherche
    Me.cmdSearch.Enabled = False
End Sub

Private Sub Recherche()
    Dim strSqlTemp As String

    If Me.lstGenus.ListIndex = -1 Then Exit Sub

    strSqlTemp = "SELECT DiatomsWithResults.CDDTaxonFullName,
DiatomsWithResults.ID_Diatom " & _
        "FROM DiatomsWithResults " & _
        "WHERE DiatomsWithResults.Genus "

    '----- Genus -----
    strSqlTemp = strSqlTemp & SqlWhere_Genus()
    '-----


If Me.lstSpecies.ListIndex <> -1 Then
    '----- Species -----
    strSqlTemp = strSqlTemp & " AND DiatomsWithResults.Species "
    strSqlTemp = strSqlTemp & SqlWhere_Species()
    '
```

```

End If

If Me.IstThirdTxName.ListIndex <> -1 Then
  '----- Third Taxon Name -----
  strSqlTemp = strSqlTemp & " AND DiatomsWithResults.Third_Taxon_Name "
  strSqlTemp = strSqlTemp & SqlWhere_ThirdTxName()
  '----- Variety -----
  strSqlTemp = strSqlTemp & " AND DiatomsWithResults.Variety "
  strSqlTemp = strSqlTemp & SqlWhere_Variety()
  '----- Form -----
  strSqlTemp = strSqlTemp & " AND DiatomsWithResults.Form "
  strSqlTemp = strSqlTemp & SqlWhere_Form()
End If

If Me.IstFourthTxName.ListIndex <> -1 Then
  '----- FourthTxName -----
  strSqlTemp = strSqlTemp & " AND DiatomsWithResults.Fourth_Taxon_Name "
  strSqlTemp = strSqlTemp & SqlWhere_FourthTxName()
  '----- Form -----
  strSqlTemp = strSqlTemp & " ORDER BY DiatomsWithResults.CDDTaxonFullName;"

  'Me.IstDiatoms.RowStyleType = "Table/Requête"
  Me.IstDiatoms.RowStyleType = "Table/Query"
  Me.IstDiatoms.RowStyle = strSqlTemp

End Sub

Private Function SqlWhere_Genus() As String
  Dim strCritereGenus As String
  If (Format(Me.IstGenus.Value) = vbNullString) Then
    strCritereGenus = " IS NULL "
  Else
    strCritereGenus = " = " & Me.IstGenus.Value & " "
  End If
  SqlWhere_Genus = strCritereGenus

```

```
End Function

Private Function SqlWhere_Species() As String
    Dim strCritereSpecies As String
    If (Format(Me.IstSpecies.Value) = vbNullString) Then
        strCritereSpecies = " IS NULL "
    Else
        strCritereSpecies = " = "" & Me.IstSpecies.Value & ""
    End If
    SqlWhere_Species = strCritereSpecies
End Function

Private Function SqlWhere_ThirdTxName() As String
    Dim strThird As String
    strThird = Format(Me.IstThirdTxName.Column(0, Me.IstThirdTxName.ListIndex))

    Dim strCritereThird As String
    If strThird = vbNullString Then
        strCritereThird = " IS NULL "
    Else
        strCritereThird = " = "" & strThird & ""
    End If
    SqlWhere_ThirdTxName = strCritereThird
End Function

Private Function SqlWhere_Variety() As String
    Dim strVar As String
    strVar = Format(Me.IstThirdTxName.Column(1, Me.IstThirdTxName.ListIndex))

    Dim strCritereVariety As String
    If strVar = vbNullString Then
        strCritereVariety = " IS NULL "
    Else
        strCritereVariety = " = "" & strVar & ""
    End If
    SqlWhere_Variety = strCritereVariety
End Function

Private Function SqlWhere_FourthTxName() As String
    Dim strFourthTxName As String
    strFourthTxName = Format(Me.IstFourthTxName.Column(0,
Me.IstFourthTxName.ListIndex))

    Dim strCritereFourthTxName As String
    If strFourthTxName = vbNullString Then
```

```
strCritereFourthTxName = " IS NULL "
Else
    strCritereFourthTxName = " = "" & strFourthTxName & ""
End If
SqlWhere_FourthTxName = strCritereFourthTxName
End Function

Private Function SqlWhere_Form() As String
Dim strForm As String
strForm = Format(Me.IstFourthTxName.Column(1, Me.IstFourthTxName.ListIndex))

Dim strCritereForm As String
If strForm = vbNullString Then
    strCritereForm = " IS NULL "
Else
    strCritereForm = " = "" & strForm & ""
End If
SqlWhere_Form = strCritereForm
End Function
```

Objets**Section: Détail****Section: EntêteFormulaire****Section: PiedFormulaire****Bouton de commande: cmdClose****Bouton de commande: cmdFirstAuthorDetails****Zone de texte: DOI_Name****Étiquette: DOI_Name_Étiquette****Étiquette: Étiquette18****Étiquette: Étiquette25****Sous-formulaire/Sous-état: Fille21****Sous-formulaire/Sous-état: Fille22****Zone de texte: ISBN****Étiquette: ISBN_Étiquette****Zone de texte: ISSN****Étiquette: ISSN_Étiquette****Zone de texte: PubFullName****Étiquette: PubFullName_Étiquette****Code**

```
VERSION 1.0 CLASS
BEGIN
    MultiUse = -1 'True
END
Attribute VB_Name = "Form_F_Publication"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = True
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Compare Database
Option Explicit
```

```
Private Sub cmdFirstAuthorDetails_Click()
    DoCmd.OpenForm "r_PEOPLE"
    Call Form_r_People.Search(Me!ID_People)
End Sub
```

Objets**Section: Détail****Section: EntêteFormulaire****Section: PiedFormulaire****Zone de texte: Area****Étiquette: Area_Etiquette****Rectangle: Boîte54****Zone de liste déroulante: cboCountryList****Zone de liste déroulante: cboLakeList****Zone de liste déroulante: cboResearcherList****Zone de liste déroulante: cboStateList****Zone de texte: Circumference****Étiquette: Circumference_Etiquette****Bouton de commande: cmdClose****Bouton de commande: cmdSearch****CtlTab49****Zone de texte: Depth_max****Étiquette: Depth_max_Etiquette****Étiquette: Etiquette24****Étiquette: Etiquette26****Étiquette: Etiquette28****Étiquette: Etiquette30****Étiquette: Etiquette32****Étiquette: Etiquette35****Étiquette: Etiquette72****Zone de texte: Lat_DD****Étiquette: Lat_DD_Etiquette****Étiquette: IbiNoResultDiatomGraph**

Étiquette: lbINoResultDiatomList

Étiquette: lbINoResultLOI

Étiquette: lbINoResultWaterChem

Zone de texte: Long_DD

Étiquette: Long_DD_Étiquette

Sous-formulaire/Sous-état: sf_Logo

Sous-formulaire/Sous-état: sfCores

Sous-formulaire/Sous-état: sfrm_Pubs

Sous-formulaire/Sous-état: sfrmDiatomGraph

Sous-formulaire/Sous-état: sfrmDiatomList

Sous-formulaire/Sous-état: sfrmLOI

Sous-formulaire/Sous-état: sfrmWaterChem

Sous-formulaire/Sous-état: sfWater

tabCore

tabCoreWater

tabDiatomGraph

tabDiatomList

tabLOI

tabWater

tabWaterChem

Zone de texte: txtLinkCoreResult

Zone de texte: txtLinkWaterResult

Code

```
VERSION 1.0 CLASS
BEGIN
    MultiUse = -1  'True
END
Attribute VB_Name = "Form_F_Result_by_Lake"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = True
Attribute VB_PredeclaredId = True
```

Attribute VB_Exposed = False

Option Compare Database

Option Explicit

```
Private Sub cboCountryList_AfterUpdate()
```

' Change la source de données pour la liste des états selon l'item sélectionné

```
cboStateList.RowSource = "SELECT StateList.ID_Country, StateList.State FROM StateList  
WHERE StateList.ID_Country = " & Me.cboCountryList.Value  
cboStateList.Requery ' redemande les infos de la base de données.
```

cboStateList.Value = cboStateList.ItemData(0) 'affiche la premiere valeur de la liste.

```
cboStateList_AfterUpdate ' simule un choix dans la liste des états pour mettre a jour celle  
des lakes.
```

' meme chose pour la liste des researchers

```
cboResearcherList.RowSource = "SELECT ResearcherList.ID_People,  
ResearcherList.Last_Name, ResearcherList.First_Name FROM ResearcherList WHERE  
ResearcherList.ID_Country = " & Me.cboCountryList.Value  
cboResearcherList.Requery  
cboResearcherList.Value = cboResearcherList.ItemData(0)
```

End Sub

```
Private Sub cboResearcherList_AfterUpdate()
```

Dim strSql As String

```
strSql = "SELECT LakeList.ID_Lake, LakeList.Lake_Code, LakeList.Lake_Name " &  
"FROM LakeList " &  
"WHERE LakeList.ID_People = " & Me.cboResearcherList.Column(0,  
Me.cboResearcherList.ListIndex) &  
" GROUP BY LakeList.ID_Lake, LakeList.Lake_Code, LakeList.Lake_Name " &  
"ORDER BY LakeList.Lake_Code "
```

cboLakeList.RowSource = strSql

cboLakeList.Requery

cboLakeList.Value = cboLakeList.ItemData(0)

cboStateList.Value = vbNullString

End Sub

```
Private Sub cboStateList_AfterUpdate()
```

Dim strSql As String

```
strSql = "SELECT LakeList.ID_Lake, LakeList.Lake_Code, LakeList.Lake_Name " &  
"FROM LakeList " &  
"WHERE LakeList.State = " & Me.cboStateList.Column(1,  
Me.cboStateList.ListIndex) & " " &
```

```

"GROUP BY LakeList.ID_Lake, LakeList.Lake_Code, LakeList.Lake_Name " & _
"ORDER BY LakeList.Lake_Code "


cboLakeList.RowSource = strSql
cboLakeList.Requery
cboLakeList.Value = cboLakeList.ItemData(0)

cboResearcherList.Value = vbNullString

End Sub

Private Sub cmdSearch_Click()
    ' Affiche seulement les informations du lake selectionne dans la recherche.
    Me.RecordSource = "select * from LakeDetails where lake_code = " &
Me.cboLakeList.Column(1, Me.cboLakeList.ListIndex) & ""
    Me.RecordSource = "select * from LakeDetails where id_Lake = " &
Me.cboLakeList.Column(0, Me.cboLakeList.ListIndex)
    Me.Requery

End Sub

Private Sub Form_Current()

    If Me.sfCores.Form.RecordsetClone.RecordCount = 0 Then Me.txtLinkCoreResult.Value =
= -1
    If Me.sfWater.Form.RecordsetClone.RecordCount = 0 Then Me.txtLinkWaterResult.Value =
= -1

    Me.tabLOI.Controls("sfrmLOI").Visible =
Me.tabLOI.Controls("sfrmLOI").Form.RecordsetClone.RecordCount > 0
    Me.tabLOI.Controls("lblNoResultLOI").Visible = Not
Me.tabLOI.Controls("sfrmLOI").Visible

    Me.tabWaterchem.Controls("sfrmWaterChem").Visible =
Me.tabWaterchem.Controls("sfrmWaterChem").Form.RecordsetClone.RecordCount > 0
    Me.tabWaterchem.Controls("lblNoResultWaterChem").Visible = Not
Me.tabWaterchem.Controls("sfrmWaterChem").Visible

    Me.tabDiatomGraph.Controls("sfrmDiatomGraph").Visible =
Me.tabDiatomGraph.Controls("sfrmDiatomGraph").Form.RecordsetClone.RecordCount > 0
    Me.tabDiatomGraph.Controls("lblNoResultDiatomGraph").Visible = Not
Me.tabDiatomGraph.Controls("sfrmDiatomGraph").Visible

    Me.tabDiatomList.Controls("sfrmDiatomList").Visible =
Me.tabDiatomList.Controls("sfrmDiatomList").Form.RecordsetClone.RecordCount > 0
    Me.tabDiatomList.Controls("lblNoResultDiatomList").Visible = Not
Me.tabDiatomList.Controls("sfrmDiatomList").Visible

End Sub

```

```
Private Sub Form_Open(Cancel As Integer)
    cboCountryList.Value = cboCountryList.ItemData(0)
    cboCountryList_AfterUpdate
    cmdSearch_Click
End Sub

Public Sub Search(IngIDLake As Long)
    ' Affiche seulement les informations du lake selectionne dans la recherche.
    Me.RecordSource = "select * from LakeDetails where ID_Lake = " & IngIDLake
    Me.Requery

    Me.cboCountryList.Value = Me!id_country
    Call cboCountryList_AfterUpdate

    Me.cboStateList.Value = Me!State
    Call cboStateList_AfterUpdate

    Me.cboLakeList.Value = Me!ID_Lake

    Me.CtlTab49.Value = 2
End Sub

Private Sub sfrm_Pubs_Enter()
    DoCmd.OpenForm "F_Publication"
End Sub
```

Objets

Section: Détail

Zone de texte: End_Depth

Étiquette: Étiquette1

Étiquette: Étiquette11

Étiquette: Étiquette5

Étiquette: Étiquette7

Étiquette: Étiquette9

Zone de texte: Full_Taxon_Name

Zone de texte: Name_Analit

Zone de texte: Result

Zone de texte: Start_Depth

Objets**Section: Détail****Section: EntêteFormulaire****Section: PiedFormulaire****Image: Image24****Image: Image26****Étiquette: Label1****Bouton de commande: Option1****Bouton de commande: Option2****Bouton de commande: Option3****Bouton de commande: Option4****Bouton de commande: Option5****Bouton de commande: Option6****Bouton de commande: Option7****Bouton de commande: Options****Étiquette: OptionLabel1****Étiquette: OptionLabel2****Étiquette: OptionLabel3****Étiquette: OptionLabel4****Étiquette: OptionLabel5****Étiquette: OptionLabel6****Étiquette: OptionLabel7****Étiquette: OptionLabel8****Image: Picture****Code****VERSION 1.0 CLASS****BEGIN****MultiUse = -1 'True**

```
END
Attribute VB_Name = "Form_Menu Général"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = True
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Compare Database
Option Explicit

Private Sub Form_Open(Cancel As Integer)
' Minimize the database window and initialize the form.

    ' Move to the switchboard page that is marked as the default.
    Me.Filter = "[ItemNumber] = 0 AND [Argument] = 'Par défaut' "
    Me.FilterOn = True

End Sub

Private Sub Form_Current()
' Update the caption and fill in the list of options.

    Me.Caption = Nz(Me![ItemText], "")
    FillOptions

End Sub

Private Sub FillOptions()
' Fill in the options for this switchboard page.

    ' The number of buttons on the form.
    Const conNumButtons = 8

    Dim con As Object
    Dim rs As Object
    Dim stSql As String
    Dim intOption As Integer

    ' Set the focus to the first button on the form,
    ' and then hide all of the buttons on the form
    ' but the first. You can't hide the field with the focus.
    Me![Option1].SetFocus
    For intOption = 2 To conNumButtons
        Me("Option" & intOption).Visible = False
        Me("OptionLabel" & intOption).Visible = False
    Next
End Sub
```

Next intOption

```
' Open the table of Switchboard Items, and find
' the first item for this Switchboard Page.
Set con = Application.CurrentProject.Connection
stSql = "SELECT * FROM [Switchboard Items]"
stSql = stSql & " WHERE [ItemNumber] > 0 AND [SwitchboardID]=" & Me! [SwitchboardID]
stSql = stSql & " ORDER BY [ItemNumber];"
Set rs = CreateObject("ADODB.Recordset")
rs.Open stSql, con, 1 ' 1 = adOpenKeyset

' If there are no options for this Switchboard Page,
' display a message. Otherwise, fill the page with the items.
If (rs.EOF) Then
    Me![OptionLabel1].Caption = "Il n'y a aucun élément pour cette page de Menu Général"
Else
    While (Not (rs.EOF))
        Me("Option" & rs![ItemNumber]).Visible = True
        Me("OptionLabel" & rs![ItemNumber]).Visible = True
        Me("OptionLabel" & rs![ItemNumber]).Caption = rs![ItemText]
        rs.MoveNext
    Wend
End If

' Close the recordset and the database.
rs.Close
Set rs = Nothing
Set con = Nothing

End Sub

Private Function HandleButtonClick(intBtn As Integer)
' This function is called when a button is clicked.
' intBtn indicates which button was clicked.

' Constants for the commands that can be executed.
Const conCmdGotoSwitchboard = 1
Const conCmdOpenFormAdd = 2
Const conCmdOpenFormBrowse = 3
Const conCmdOpenReport = 4
Const conCmdCustomizeSwitchboard = 5
Const conCmdExitApplication = 6
Const conCmdRunMacro = 7
Const conCmdRunCode = 8
```

```

Const conCmdOpenPage = 9

' An error that is special cased.

Const conErrDoCmdCancelled = 2501

Dim con As Object
Dim rs As Object
Dim stSql As String

On Error GoTo HandleButtonClick_Err

' Find the item in the Switchboard Items table
' that corresponds to the button that was clicked.

Set con = Application.CurrentProject.Connection
Set rs = CreateObject("ADODB.Recordset")
stSql = "SELECT * FROM [Switchboard Items]"
stSql = stSql & "WHERE [SwitchboardID] = " & Me![SwitchboardID] & " AND
[ItemNumber] = " & intBtn
rs.Open stSql, con, 1 ' 1 = adOpenKeyset

' If no item matches, report the error and exit the function.

If (rs.EOF) Then
    MsgBox "Une erreur s'est produite lors de la lecture de la table d'élément du Menu
général."
    rs.Close
    Set rs = Nothing
    Set con = Nothing
    Exit Function
End If

Select Case rs![Command]

' Go to another switchboard.

Case conCmdGotoSwitchboard
    Me.Filter = "[ItemNumber] = 0 AND [SwitchboardID] = " & rs![Argument]

' Open a form in Add mode.

Case conCmdOpenFormAdd
    DoCmd.OpenForm rs![Argument], , , acAdd

' Open a form.

Case conCmdOpenFormBrowse
    DoCmd.OpenForm rs![Argument]

' Open a report.

Case conCmdOpenReport

```

```
DoCmd.OpenReport rs![Argument], acPreview

' Customize the Switchboard.
Case conCmdCustomizeSwitchboard
    ' Handle the case where the Switchboard Manager
    ' is not installed (e.g. Minimal Install).
    On Error Resume Next
    Application.Run "ACWZMAIN.sbm_Entry"
    If (Err <> 0) Then MsgBox "Commande non disponible."
    On Error GoTo 0
    ' Update the form.
    Me.Filter = "[Item Number] = 0 AND [Argument] = 'Par défaut' "
    Me.Caption = Nz(Me![ItemText], "")
    FillOptions

' Exit the application.
Case conCmdExitApplication
    CloseCurrentDatabase

' Run a macro.
Case conCmdRunMacro
    DoCmd.RunMacro rs![Argument]

' Run code.
Case conCmdRunCode
    Application.Run rs![Argument]

' Open a Data Access Page
Case conCmdOpenPage
    DoCmd.OpenDataAccessPage rs![Argument]

' Any other command is unrecognized.
Case Else
    MsgBox "Option inconnue."

End Select

' Close the recordset and the database.
rs.Close

HandleButtonClick_Exit:
On Error Resume Next
Set rs = Nothing
Set con = Nothing
```

Exit Function

```
HandleButtonClick_Err:  
    ' If the action was cancelled by the user for  
    ' some reason, don't display an error message.  
    ' Instead, resume on the next line.  
    If (Err = conErrDoCmdCancelled) Then  
        Resume Next  
    Else  
        MsgBox "Une erreur s'est produite lors de l'exécution de la commande.", vbCritical  
        Resume HandleButtonClick_Exit  
    End If  
  
End Function
```

Objets

Section: Détail

Section: EntêteFormulaire

Section: PiedFormulaire

Rectangle: Boîte91

Rectangle: Boîte95

Zone de liste déroulante: cboPeople

Bouton de commande: cmdClose

CtlTab30

Zone de texte: Email

Étiquette: Email_Etiquette

Étiquette: Étiquette24

Étiquette: Étiquette26

Étiquette: Étiquette40

Étiquette: Étiquette44

Étiquette: Étiquette46

Étiquette: Étiquette48

Étiquette: Étiquette49

Étiquette: Étiquette50

Étiquette: Étiquette52

Étiquette: Étiquette54

Étiquette: Étiquette56

Étiquette: Étiquette83

Étiquette: Étiquette85

Étiquette: Étiquette87

Étiquette: Étiquette93

Étiquette: Étiquette94

Sous-formulaire/Sous-état: F_Visit

Zone de texte: Fax_Number

Étiquette: Fax_Number_Étiquette

Zone de texte: First_Name

Étiquette: First_Name_Étiquette

Zone de texte: Last_Name

Étiquette: Last_Name_Étiquette

Zone de texte: Lib_UNIV.Full_Address

Étiquette: List_Étiquette

Page81

Zone de texte: PEOPLE.Full_Address

Zone de texte: Phone

Étiquette: Phone_Étiquette

Zone de texte: Postal_Code

Étiquette: Postal_Code_Étiquette

Sous-formulaire/Sous-état: r_Affil

Sous-formulaire/Sous-état: sf_Logo

Sous-formulaire/Sous-état: sf_PeoplePub

Étiquette: sf_PeoplePubÉtiquette

tabPeopleAffill

tabPeopleDetails

tabPeopleVisit

Zone de texte: Texte43

Zone de texte: Texte45

Zone de texte: Texte47

Zone de texte: Texte51

Zone de texte: Texte53

Zone de texte: Texte55

Zone de texte: Texte82

Zone de texte: Texte84

Zone de texte: Texte86

Zone de texte: Titre

Étiquette: Titre_Etiquette

Zone de texte: txtLinkPeople_1

Zone de texte: Univ_Name

Étiquette: Univ_Name_Etiquette

Zone de texte: Webpage

Étiquette: Webpage_Etiquette

Code

```
VERSION 1.0 CLASS
BEGIN
    MultiUse = -1  'True
END
Attribute VB_Name = "Form_r_People"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = True
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
Option Compare Database

Private Sub cboPeople_AfterUpdate()
    Search (Me.cboPeople.Column(0))
End Sub

Public Sub Search(idPeople As Long)
    Dim rs As dao.Recordset
    Dim bd As dao.Database

    Set bd = CurrentDb
    Set rs = Me.RecordsetClone
    rs.FindFirst ("ID_People = " & idPeople)
    If Not rs.NoMatch Then
        If Me.FilterOn Then Me.FilterOn = False
        Me.Bookmark = rs.Bookmark
        Me.cboPeople.Value = idPeople
    End If

    rs.Close
End Sub
```

Set rs = Nothing

Set bd = Nothing

End Sub

Objets

Section: Détail

Zone de texte: Core_Code

Zone de texte: End_Depth

Étiquette: Étiquette1

Étiquette: Étiquette11

Étiquette: Étiquette3

Étiquette: Étiquette5

Étiquette: Étiquette7

Étiquette: Étiquette9

Zone de texte: Full_Taxon_Name

Zone de texte: Name_Analit

Zone de texte: Result

Zone de texte: Start_Depth

Code

```
VERSION 1.0 CLASS
BEGIN
    MultiUse = -1 'True
END
Attribute VB_Name = "Form_TableauCroisDyn TEST_1"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = True
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = False
```


Annexe D
Code de programmation macro (MS Excel)

Macro Excel

```

Sub MetDonneesEnForme2Modifiée()
    ' Before running this, make sure you have a summary table with column headers.
    ' The output table will have three columns.

    Dim SummaryTable As Range, OutputRange As Range
    Dim OutRow As Long
    Dim r As Long, c As Long

    On Error Resume Next
    Set SummaryTable = ActiveCell.CurrentRegion
    If SummaryTable.Count = 1 Or SummaryTable.Rows.Count < 3 Then
        MsgBox "Select a cell within the summary table.", vbCritical
        Exit Sub
    End If
    SummaryTable.Select
    Set OutputRange = Application.InputBox(prompt:="Sélectionnez une cellule de départ pour la forme2.", Type:=8)
    ' Convert the range
    OutRow = 2
    Application.ScreenUpdating = False
    OutputRange.Range("A1:E1") = Array("Lake_Code", "Start_Level", "End_Level",
    "Diatom", "Result")

    ' Boucle à partir de la 2ième ligne et pour toutes les lignes
    For r = 2 To SummaryTable.Rows.Count

        ' Boucle à partir de la 4ième colonne et pour toutes les colonnes
        For c = 4 To SummaryTable.Columns.Count
            OutputRange.Cells(OutRow, 1) = SummaryTable.Cells(r, 1) ' Colonne Lake_Code
            OutputRange.Cells(OutRow, 2) = SummaryTable.Cells(r, 2) ' Colonne Start_Level
            OutputRange.Cells(OutRow, 3) = SummaryTable.Cells(r, 3) ' Colonne End_Level
            OutputRange.Cells(OutRow, 4) = SummaryTable.Cells(1, c) ' Colonne Diatom
            OutputRange.Cells(OutRow, 5) = SummaryTable.Cells(r, c) ' Colonne Result
            OutputRange.Cells(OutRow, 5).NumberFormat = SummaryTable.Cells(r,
c).NumberFormat
            OutRow = OutRow + 1
        Next c
    Next r
End Sub

```


Annexe E
Codes de programmation VBA (MS Access)

Form_F_Abund_by_Diatom

Option Compare Database

Option Explicit

```
Private Sub cmdGoToWaterResult_Click()
```

```
Dim IngLakeID As Long
```

```
IngLakeID = Me.sfrmResultDiatoms!ID_Lake
```

```
DoCmd.OpenForm "F_Result_by_Lake"
```

```
Form_F_Result_by_Lake.Search (IngLakeID)
```

```
End Sub
```

Form_SF_Core per Lake

Option Compare Database

Option Explicit

```
Private Sub Core_Code_DblClick(Cancel As Integer)
```

```
DoCmd.OpenForm "F_Tab_Result per level", acNormal, , "[ID_CORE] = " & Me!ID_Core  
End Sub
```

```
Private Sub Form_Current()
```

```
Me.Parent.Form.txtLinkCoreResult.Value = Nz(Me.ID_Core, -1)  
End Sub
```

Form_sf_Pubs_by_Visit**Option Compare Database****Option Explicit****Private Sub PubAuthor_Click()****DoCmd.OpenForm "F_Publication", acNormal, , "ID_PUB = " & Me!ID_Pub
End Sub**

Form_SF_WaterByLake

Option Compare Database

Option Explicit

Private Sub Form_Current()

Me.Parent.Form.txtLinkWaterResult.Value = Nz(Me.ID_Water_Chem, -1)

End Sub

Form_SF_Core per Lake

Option Compare Database

Option Explicit

```
Private Sub Core_Code_DblClick(Cancel As Integer)
```

```
DoCmd.OpenForm "F_Tab_Result per level", acNormal, , "[ID_CORE] = " & Me!ID_Core
End Sub
```

ImportationExcel

Option Compare Database

Option Explicit

```
Public Function CreeCoreLevel(lake_name As String, startdepth As Double, enddepth As Double)
```

```
As Double
```

```
Dim noseq As Double
```

```
Dim bd As DAO.Database
```

```
Dim rsCore As DAO.Recordset
```

```
Dim rsCoreLevel As DAO.Recordset
```

```
Set bd = CurrentDb
```

```
Set rsCoreLevel = bd.OpenRecordset("CORE_LEVEL")
```

```
Set rsCore = bd.OpenRecordset("SELECT ID_CORE FROM CORE WHERE CORE_CODE = '" &  
lake_name & "'")
```

```
With rsCoreLevel
```

```
    .AddNew
```

```
    !ID_Core = rsCore!ID_Core
```

```
    !start_depth = startdepth
```

```
    !end_depth = enddepth
```

```
    .Update
```

```
noseq = !id_level
```

```
End With
```

```
CreeCoreLevel = noseq
```

```
rsCore.Close
```

```
rsCoreLevel.Clone
```

```
bd.Close
```

```
Set rsCore = Nothing  
Set rsCoreLevel = Nothing  
Set bd = Nothing
```

```
End Function
```

Supress_doubloons**Option Compare Database**

```
Sub DeleteDuplicateRecords(strTableName As String)
' Deletes exact duplicates from the specified table.
' No user confirmation is required. Use with caution.
Dim rst As DAO.Recordset
Dim rst2 As DAO.Recordset
Dim tdf As DAO.TableDef
Dim fld As DAO.Field
Dim strSQL As String
Dim varBookmark As Variant

Set tdf = DBEngine(0)(0).TableDefs(strTableName)
strSQL = "SELECT * FROM " & strTableName & " ORDER BY "
' Build a sort string to make sure duplicate records are
' adjacent. Can't sort on OLE or Memo fields,though.
For Each fld In tdf.Fields

If (fld.Type <> dbMemo) And (fld.Type <> dbLongBinary) Then
strSQL = strSQL & fld.Name & ", "

End If
Next fld
' Remove the extra comma and space from the SQL
strSQL = Left(strSQL, Len(strSQL) -2)
Set tdf = Nothing

Set rst = CurrentDb.OpenRecordset(strSQL)
Set rst2 = rst.Clone
rst.MoveNext
Do Until rst.EOF
```

```
varBookmark = rst.Bookmark  
For Each fld In rst.Fields  
If fld.Value <> rst2.Fields(fld.Name).Value Then  
GoTo NextRecord
```

```
End If  
Next fld  
rst.Delete  
GoTo SkipBookmark
```

```
NextRecord:  
rst2.Bookmark = varBookmark  
SkipBookmark:  
rst.MoveNext  
Loop  
End Sub
```

