

BADJAM

LA BANQUE DES DONNEES JAMESIENNES A ACCES MULTIPLES

Didier Bussy
Responsable de la banque des données écologiques
Société de développement de la Baie James
800 est, boul. de Maisonneuve
Bureau 2200
Montréal, Québec
H2L 4M6

RESUME

L'un des mandats de la Société de développement de la Baie James (SDBJ) est d'étudier les effets du projet de la Baie James sur l'environnement et de les optimiser. A cet effet, la SDBJ a mis sur pied une banque des données de la cartographie écologique du territoire de la Baie James. Dans cet article, on décrit les données mises en banque, puis le système de gestion de la banque de données et ses capacités pour l'interrogation. Enfin, une dernière partie est consacrée aux développements en cours qui mèneraient, en particulier, à l'implantation d'un système de cartographie automatique.

ABSTRACT

One of the mandates of the Société de développement de la Baie James is to study the impact of the James Bay project on the environment and to make this impact as favourable as possible. For this purpose the SDBJ has established a bank of data relating to the ecological geography of the territory. This article describes the data, the management of the data bank, and the methods of access to the data. The last part of the article deals with future developments and especially with the creation of a system capable of generating maps by computer.

INTRODUCTION

Le projet hydro-électrique de la baie James: le projet du siècle a-t-on déjà dit! C'est aussi celui de l'aménagement et du développement d'un territoire immense encore vierge, 350 000 kilomètres carrés (135 000 milles carrés) près d'un cinquième de toute l'étendue du Québec. C'est-à-dire plus du double de la superficie de l'Angleterre, plus de la moitié de celle de la France. Les limites du sud du territoire se situent à 480 km (300 mi) au nord de Montréal. Quand au nord du territoire, il est à quelque 1 120 km (700 mi) de Montréal (fig. 1).

La Société de développement de la Baie James de concert et en vertu d'une entente avec Environnement Canada a mis sur pied un vaste plan de collecte de données biophysiques. Entre autres, une cartographie écologique exhaustive du territoire a été réalisée. La Société répond ainsi au souci d'étudier d'une manière globale les effets du projet de la Baie James sur l'environnement et de les optimiser.

Les problèmes que nous avons rencontrés dès le début du projet de banque de données étaient de différents ordres. En ce qui concerne les données, il y avait tout d'abord un problème de quantité. En effet, des inventaires de toutes sortes, allant de celui du castor à celui des sites archéologiques en passant par la cartographie écologique qu'on mentionnait tout à l'heure, produisait une quantité énorme de données.

Le second handicap qui s'est présenté était celui de la diversité des supports des données. En effet, les données résidaient sur des cartes géographiques, des fiches manuscrites, des photographies aériennes, des rapports, etc...

Le troisième sujet de préoccupation concernait l'identification des besoins des utilisateurs possibles de ces données.

Après consultations et analyses, la décision de construire une banque de données informatisée s'est imposée. Le principal argument pour appuyer cette décision tenait à la quantité de données à traiter.

Nous avons opté pour mettre sur pied une banque de données de la cartographie écologique qui couvre le territoire de la Baie James. La cartographie écologique constitue un groupe homogène de données qui peut produire des informations tant aux spécialistes de la protection de l'environnement qu'à ceux de l'aménagement du territoire.

1. LES DONNEES

1.1 La carte écologique

La méthodologie utilisée pour dresser la carte écologique est la suivante:

- i) étude des photographies aériennes au 1:60 000
- ii) cartographie préliminaire
- iii) vérifications sur le terrain
- iv) cartographie finale

La carte écologique est constituée de 123 feuillets au 1:125 000. Chaque feuillet est identifié selon le quadrillage de Energie, Mines et Ressources du Canada. Par exemple: 33 FNE est un identificateur de feuillet (fig. 2). Chaque feuillet de la carte montre:

- i) les contours des écosystèmes terrestres et leurs définitions;
- ii) les contours des écosystèmes aquatiques et leurs définitions;

On notera au passage que les contours de ces 2 sortes d'écosystèmes ne se confondent que rarement.

De plus, en se basant sur des considérations relatives au couvert végétal, on trouve sur cette carte le tracé des 19 régions écologiques du territoire. Par exemple la région du lac Mistassini est notée MI.

1.1.1 Les écosystèmes terrestres. Au niveau global de la carte écologique, les écosystèmes terrestres sont définis par 8 descripteurs soit (fig. 3):

- i) la région
- ii) le relief
- iii) l'épaisseur des matériaux meubles
- iv) la nature et l'origine des matériaux géologiques de surface dominants
- v) morphologie des matériaux de surface dominants
- vi) nature et origine pour les matériaux sous dominants
- vii) morphologie pour les matériaux sous dominants
- viii) numéro d'ordre de l'écosystème terrestre

Chaque descripteur varie à l'intérieur d'une fourchette de valeur. Par exemple, le relief -selon le milieu dans lequel on se trouve- prendra les valeurs ordonnées F, U, R, H, M; F représentant un relief plat et M un relief montagneux.

1.1.2 Les écosystèmes aquatiques. Ils sont, au niveau global de la carte écologique, définis par 3 descripteurs soit (fig. 3):

- i) la catégorie d'écosystème aquatique;
- ii) l'abondance des ruisseaux;
- iii) l'abondance des Wetlands.

1.1.3 Les "cellules". La "cellule" est une entité spatiale sans valeur scientifique. C'est un subterfuge pour unifier la description d'un espace de la carte écologique (fig. 4).

La "cellule" a été créée par nous, gens du traitement de l'information, pour établir une relation univoque entre une description du milieu (chaîne unique de valeurs des descripteurs que nous avons vus plus haut) et une aire géographique unique du territoire.

Les "cellules", du point de vue spatial, sont les parties communes entre tout écosystème terrestre et les écosystèmes aquatiques qui le couvrent. Par exemple, si 3 écosystèmes aquatiques se trouvent à empiéter sur 1 écosystème terrestre donné, nous serons en présence de 3 cellules.

Pour continuer dans la même veine, nous avons dessiné - par superposition sur la carte écologique - une carte dite "carte de numérisation". Cette carte a pour fonction de recenser, à l'aide d'un identificateur de feuillet et d'un numéro, toutes les cellules du territoire. Le nom "numérisation" vient du fait que nous avons attribué un numéro différent à chaque "cellule" à l'intérieur d'un même feuillet. Le même numéro peut être donné à une autre "cellule" dans un autre feuillet ce qui le différencie alors, c'est l'identificateur de feuillet.

La "carte de numérisation" est constituée d'autant de feuillets qu'il y en a dans la carte écologique telle que nous l'avons vu. Chaque "feuillet de numérisation" contient environ 400 "cellules" (numérotées de 1 à 400). Pour l'ensemble du territoire, il y a donc aux alentours de 40 000 cellules.

Le travail de "numérisation" d'un feuillet étant accompli, les données qu'on possède sur chaque cellule sont transposées sur un formulaire qu'on appelle une "fiche-cellule". Cette "fiche-cellule" contient alors les descriptions pour:

- i) l'identificateur du feuillet dans lequel la "cellule" se trouve
- ii) le numéro de "cellule",
- iii) les 8 descripteurs de l'écosystème terrestre auquel la "cellule" appartient,

iv) les 3 descripteurs de l'écosystème aquatique auquel la cellule appartient; soit au total 13 descriptions par "fiche-cellule" pour chacune des 40 000 "cellules". Nous sommes donc en présence d'un fichier de 40 000 "fiches-cellules". Celles-ci ont été perforées sur cartes mécanographiques.

1.2 Les "fiches-types"

Pour chaque écosystème terrestre de la carte écologique, une "fiche-type" a été produite lors de la cartographie écologique. Cette "fiche-type" est un complément d'informations à la carte écologique.

La "fiche-type" contient:

- a) un complément d'informations pour les écosystèmes aquatiques qui couvrent l'écosystème terrestre que la "fiche-type" veut décrire,
- b) un complément d'informations pour l'écosystème terrestre que la "fiche-type" décrit. Ce complément est sous la forme de descriptions des types écologiques,
- c) les mêmes informations que la "fiche-cellule" sauf pour l'identificateur de carte et le numéro de cellule.

En ce qui concerne le complément d'informations pour les écosystèmes aquatiques, on trouve dans la "fiche-type", et pour chacun de ces derniers, les éléments suivants:

- a) caractéristiques morphométriques des plans d'eau
 - découpage du périmètre aquatique pour les lacs et forme du cours d'eau pour les rivières;
 - pente de la berme pour les lacs et présence de rapides pour les sources;
 - pente des rives;
 - système de drainage et profondeur;
- b) la nature des matériaux géologiques de surface des rives.

Ce complément d'informations pour les écosystèmes aquatiques s'exprime en 6 descripteurs.

BADJAM

En ce qui concerne le complément d'informations pour l'écosystème terrestre que la "fiche-type" décrit, on y trouve les définitions des types écologiques qui composent l'écosystème.

Le type écologique se définit comme suit: "une portion de territoire caractérisée par une combinaison relativement uniforme du sol et de la chronoséquence végétale" (M. Jurdant, J.L. Bélair, V. Gerardin et J.P. Ducruc, 1977).

L'écosystème terrestre décrit par la "fiche-type" est un ensemble d'unités homogènes ou types écologiques qui sont identifiés dans la "fiche-type" mais non cartographiées. A l'expérience, nous avons pu déterminer qu'il y a une moyenne de 6 types écologiques (avec un maximum de 10 et un minimum de 1) par écosystème terrestre.

Un type écologique est identifié dans la "fiche-type" par:

- i) 2 ou 3 lettres désignant la série de sol;
- ii) 1 ou 2 chiffres désignant la classe de sol;
- iii) un multiple de 5 compris entre 0 et 100 désignant l'occupation du sol qu'a relativement ce type par rapport à la surface de l'écosystème terrestre.

Un type écologique s'exprime donc en 3 descripteurs.

Pour résumer, la "fiche-type" contient donc:

- a) un complément d'informations pour chacun des écosystèmes aquatiques couvrant l'écosystème terrestre décrit. Et pour chacun, ce complément s'exprime en 6 descripteurs;
- b) un complément d'informations pour l'écosystème terrestre décrit. Ce complément à la forme de types écologiques. Chaque type écologique est identifié par 3 descripteurs
- c) au moins les mêmes informations que la "fiche-cellule" sauf pour l'identificateur de carte et le numéro de cellule.

Il est à noter qu'un écosystème terrestre peut être récurrent à travers le territoire de la Baie James. En d'autres mots, on peut retrouver le même écosystème terrestre à l'intérieur d'un même feuillet comme dans plusieurs feuillets. Et, comme il n'est défini qu'une fois -la première fois qu'on le rencontre dans le territoire- cela implique qu'il y a

"seulement" 11 500 "fiches-types". Si on compare ce nombre avec celui de 40 000 "cellules", on comprend qu'une "fiche-type" va fournir le complément d'informations pour plus d'une "cellule".

2. LE FICHIER DE BANQUE DE DONNEES

C'est le produit de la fusion du fichier des "fiches-cellules" et du fichier des "fiches-types". Comme nous l'avons vu une "fiche-type" contient, en particulier, les mêmes informations qu'une "fiche-cellule" moins l'identificateur de carte et le numéro de cellule.

Une solution simple pour réaliser la fusion consiste à compléter une "fiche-cellule" par le complément contenu dans la "fiche-type" correspondante, en créant ainsi une fiche consolidée.

On peut ajouter, bout à bout, aux descriptions d'une "cellule", les données qui se rapportent à cette dernière et qui sont contenues dans la "fiche-type".

On obtient pour les 40 000 "cellules", 40 000 fiches consolidées exprimées en un maximum de 49 descripteurs.

Selon cette solution, le fichier de banque de données atteint le volume de 1 960 000 positions de mémoire.

Cependant, l'interrogation d'une banque ainsi structurée est très laborieuse et non appropriée au genre de requêtes que les utilisateurs désirent faire.

Par exemple, si on veut extraire toute l'information concernant une "cellule" en particulier, il faut énumérer - un à un - les 49 descripteurs de la fiche consolidée.

De plus, étant donné qu'il y a une possibilité d'avoir 10 types écologiques dans une même fiche consolidée, il faut indiquer les descripteurs relatifs à ces types.

Ceci alourdit considérablement la manipulation. Autant d'opérations entraîne fatalement des erreurs de formulation.

Nous avons donc abandonné cette solution au profit de celle dont la description suit.

Au lieu de supposer que les types écologiques sont des compléments à la "cellule", nous avons considéré que la "cellule" est un complément à chacun des types écologiques.

BADJAM

Ainsi, pour une "fiche-type" contenant 5 types écologiques, nous avons 5 fois la même "fiche-cellule" qui complète chacun des types.

La fiche consolidée obtenue est beaucoup plus courte. Elle est maintenant exprimée en 22 descripteurs (au lieu de 49 dans la première solution).

Une première objection vient immédiatement à l'esprit. Cette structure est redondante.

En effet, avec une moyenne de 6 types écologiques par écosystème terrestre, le fichier de banque de données possède 240 000 fiches consolidées (40 000 "cellules multipliées par 6).

Chaque fiche consolidée ayant 22 descriptions cela fait 5 180 000 positions de mémoire.

Même si un gros ordinateur peut supporter un tel volume, on sera rarement obligé de traiter le fichier entier en accès direct, comme cela sera explicité un peu plus tard.

Une deuxième objection concerne l'interrogation. Va-t-on extraire l'information avec toute sa redondance? Le problème est écarté lorsqu'on utilise un système de banque de données qui élimine systématiquement les redondances indésirables à l'interrogation.

Nous utilisons le système EXIR qui va être décrit tout de suite.

3. LE SYSTEME DE GESTION DE LA BANQUE DE DONNEES (SGBD)

Exir a été développé au laboratoire de taximétrie de l'Université du Colorado pour supporter de grosses banques données dans le domaine de l'environnement. Il a été utilisé aux U.S.A. ces 8 dernières années pour régler une multitude de problèmes de traitement de données.

Les données sont stockées par Exir dans une forme compactée consistant en des fonctions caractéristiques de base, de concert avec un dictionnaire. Cette méthode de compaction a l'avantage de réduire considérablement le fichier de banque de données dont il a été question déjà.

BADJAM

Nous avons donc importé EXIR à Montréal. Pour des raisons techniques de configuration d'ordinateur, il a fallu en réduire à deux reprises, les dimensions. Il existe maintenant, dans une version compactée (par 4) mais tout aussi efficace que la version originale.

Sans nous étendre sur les multiples possibilités de ce système mentionnons en tout de même quelques caractéristiques.

En ce qui a trait à la construction et à la gestion d'une banque de données, le système permet notamment, de définir la structure de la banque de données à l'aide de descripteurs (dans notre cas 22).

On définit alors le nom des descripteurs, l'ordre dans lequel ils apparaissent dans les fiches consolidées et le genre de chacun des descripteurs. Par exemple, s'il s'agit de descriptions appartenant à un ensemble ordonné on indiquera que c'est un descripteur de genre "ordre".

Les données prêtes à être structurées peuvent résider sur cartes perforées ou sur fichier (codé en binaire). De plus, elles peuvent être en format fixe ou en format libre. Ceci démontre bien la facilité avec laquelle on peut construire une banque de données à partir d'à peu près n'importe quel fichier.

Une fois, une banque construite on peut, éliminer et corriger -en accès direct- des descriptions et des fiches. De plus, on peut définir d'autres descripteurs temporairement ou de façon permanente. Toutes ces manipulations se font par des commandes simples utilisant des équations booléennes.

4. L'UTILISATION DU SYSTEME

Considérant l'immensité du territoire de la Baie James, rares sont les individus intéressés à obtenir des informations portant sur l'ensemble du territoire.

A ce jour, seuls, certains aménagistes qui veulent découvrir le territoire en régions ont requis ce genre de service. Nous leur offrons la possibilité de le faire.

Cependant, la plupart du temps, les utilisateurs travaillent sur des projets précis: aménagement d'un lac, tracé d'un corridor de route ou de ligne de transport d'électricité, établissement d'une mine ou d'un camp, etc... Au niveau spatial, leur champ d'action se limite donc, à quelques feuillets de la carte écologique.

BADJAM

Pour répondre à cette dernière catégorie d'utilisateurs, nous avons mis au point une série de programmes visant à répondre plus directement aux besoins de ces derniers tout en optimisant les coûts d'utilisation. La procédure normale pour réagir face à une requête est comme suit:

- i) extraction à même le fichier de banque de données des seules données (mais de toutes) concernant la région mentionnée dans la demande spécifique. Ces données sont copiées sur disque magnétique, donc accessibles en accès direct jusqu'à la fin des opérations;
- ii) selon les besoins, un programme peut tailler à même ce fichier, pour ne conserver que les données relatives à la requête.

Ces premières opérations se font, très simplement, par télétype -en moins d'un quart d'heure (temps de la montre)- et pour un coût modeste.

Le fichier ainsi obtenu, peut -par la suite- être soumis au contrôle d'EXIR qui nous le retourne sous forme d'une banque de données interrogeable immédiatement, en mode conversationnel. Cette opération requiert, elle aussi, un temps et un coût minimes.

Il faut souligner l'extrême souplesse avec laquelle, dans un premier temps on peut cerner la zone qui nous intéresse, et dans un deuxième temps -mettre les données sous forme interrogeable.

4.1 L'interrogation en mode conversationnel

Généralement, ce mode d'interrogation intéresse des personnes qui veulent avoir accès à une donnée ou à un groupe de données pour consultation immédiate.

Par exemple, le requérant veut savoir, si tel matériau existe dans telle localité qu'il a encerclée sur une carte.

BADJAM

Après les opérations d'usage dont nous avons parlées plus haut, nous pouvons répondre à la demande en l'exprimant par une association booléenne de descripteurs.

La réponse étant reçue presque instantanément, le requérant peut modifier et/ou préciser à volonté ses questions.

4.2 L'interrogation de façon prédéfinie

Un certain nombre de requêtes comme la détermination de potentiels du milieu reviennent régulièrement. Souvent elles exigent une programmation complexe. C'est le cas des clés interprétatives de données.

Par exemple, un spécialiste du castor est intéressé à développer une méthode d'interprétation des données qui lui permet de déterminer le potentiel pour cet animal quelque soit un endroit donné du territoire.

Partant des informations qu'il a acquises sur le milieu propice au castor, le spécialiste peut -par pondération des facteurs biophysiques relevés dans la carte écologique- établir une clé interprétative qui produit le potentiel pour l'animal.

Nous avons remarqué que l'utilisateur améliore sa clé après essais successifs. Pour ne pas avoir à modifier la programmation -assez complexe- de ces clés, nous avons développé un petit compilateur booléen d'aide à la programmation.

Ainsi, l'utilisateur peut -en employant les mots du vocabulaire spécialisé- programmer et modifier sa clé à volonté.

Il existe 35 clés interprétatives pour le moment et 15 sont programmées et conservées en "programmathèque".

Les opérations d'usage de construction de banques étant exécutées, le requérant peut alors -en traitement par lots- soumettre sa clé à l'ordinateur.

Les résultats peuvent être imprimés sur un "listing" d'ordinateur ou/et être enregistrés sur un disque ou une bande magnétique pour le cas où nous voudrions faire des traitements ultérieurs.

5. DEVELOPPEMENTS EN COURS

Il y a essentiellement 2 sortes de développements en cours:

BADJAM

- 1) un logiciel de statistiques.
Il permettra aux spécialistes de faire des analyses de données. Par exemple, il serait intéressant de pouvoir faire une analyse en composantes principales sur les espèces végétales rencontrées dans telle ou telle région du territoire
- 2) un logiciel de cartographie automatique.
Ce logiciel est déjà très avancé. Il existe actuellement des programmes qui permettent l'enregistrement des contours des "cellules".

Nous avons à ce jour, enregistrés les contours des "cellules" d'environ 1/5 du territoire, sous forme de banques de coordonnées.

Il existe aussi des programmes de traçage automatique de cartes.

Ces programmes feront la liaison entre les résultats des clés interprétatives enregistrés sur disque ou bande magnétique et les banques de coordonnées.

De cette façon, on pourra présenter sous forme cartographique les résultats produits par la banque des données alphanumériques.

6. CONCLUSION

Voilà un bref aperçu du système BADJAM mis sur pied à la Société de Développement de la Baie James.

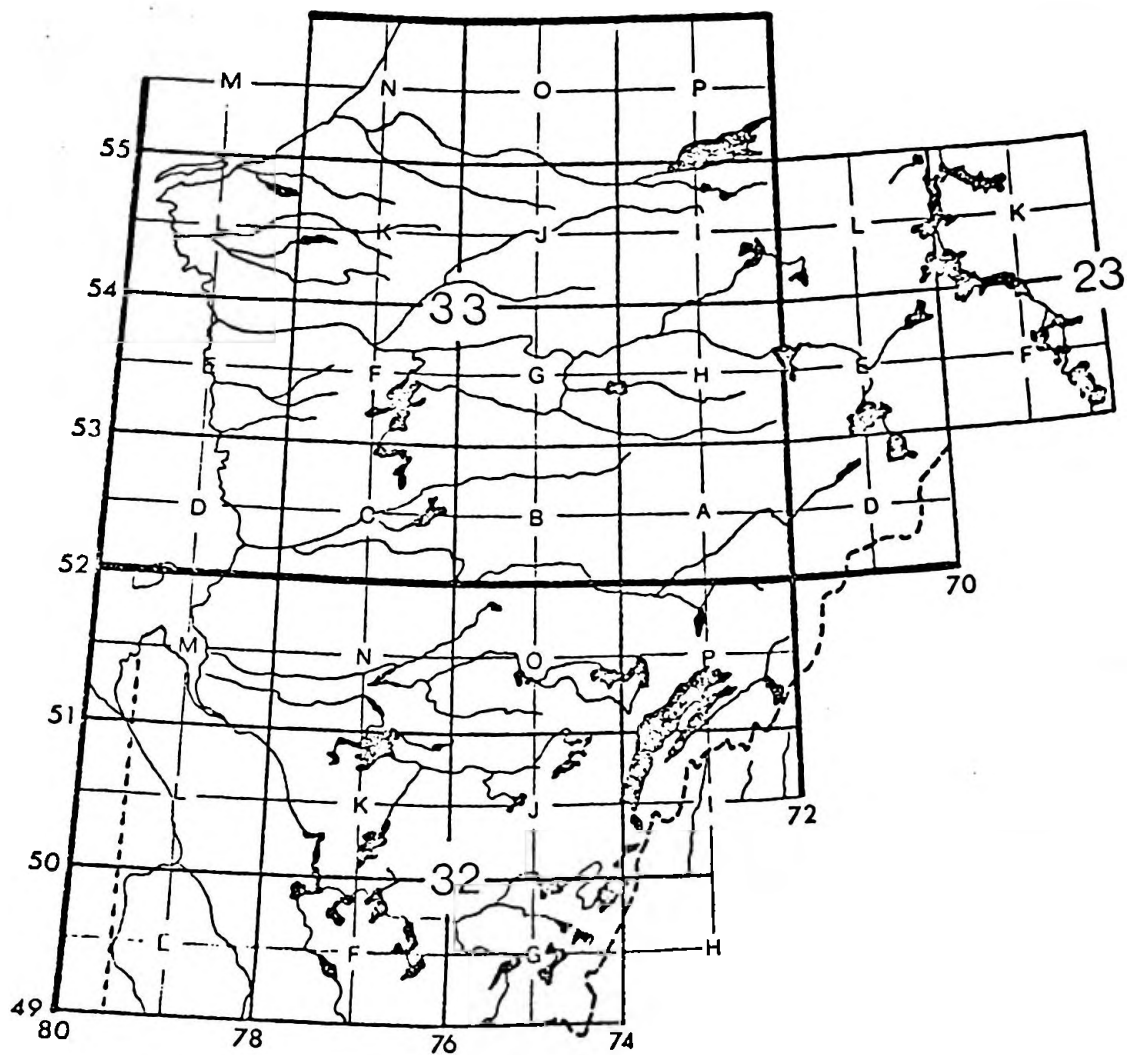
Dans cet exposé, nous avons -par respect des limites de temps qui nous sont imparties- passé sous silence, des aspects que certains auraient aimé voir développés. Par exemple, nous avons dépensé beaucoup d'énergie à mettre au point un système de validation des données. Les personnes intéressées peuvent obtenir tous les renseignements en nous contactant.

Comme nous l'avons vu, BADJAM est un système souple et efficace. Il s'inscrit comme un instrument de mise en exploitation d'un grand volume de données relevées sur le territoire de la Baie James.

Il est un outil privilégié pour veiller à la protection de l'environnement et aider au développement harmonieux des ressources du territoire.

REFERENCES BIBLIOGRAPHIQUES

- CANADA. Ministère de l'environnement. Direction générale des terres. Service des études écologiques régionales. Inventaire du capital-nature. Méthode de classification et de cartographie écologique du territoire par M. Jurdant, J.L. Bélair et V. Gérardin. (Québec) 1974.

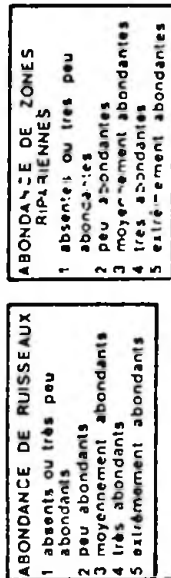
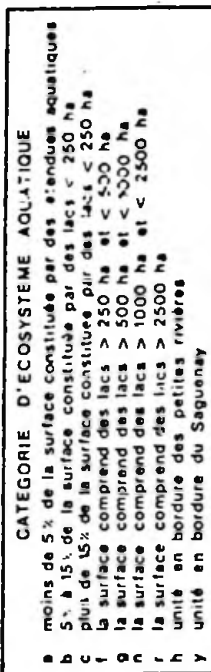


LES SYSTEMES ECOLOGIQUES

Echelle 1:125.000



Le SYSTEME ECOLOGIQUE est une portion de territoire caractérisée par un pattern propre du relief, des matériaux géologiques de surface, des sols, des chronoséquences végétales et des plans d'eau.

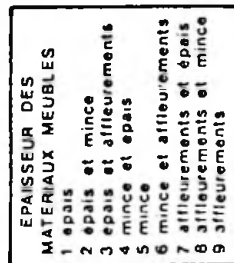
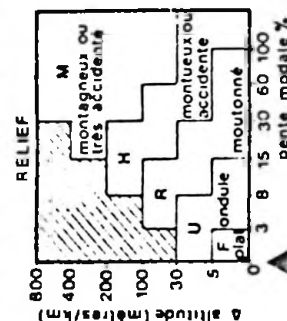


NUMERO D'ORDRE DE L'UNITE

Le numéro d'ordre rétière au rapport accompagnant la carte écologique et qui comprend la description du pattern des TYPES ECOLOGIQUES constituant l'unité.

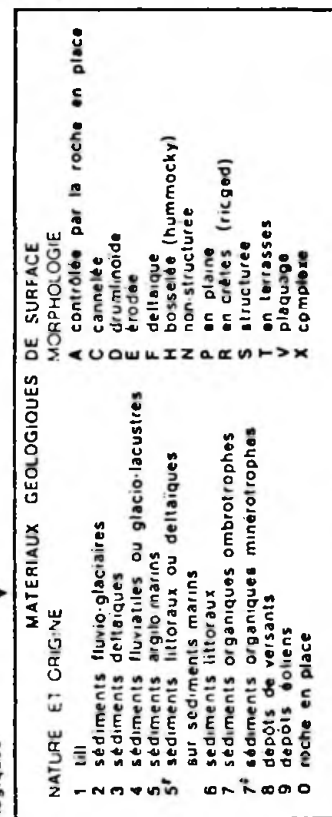
Le TYPE ECOLOGIQUE est une portion de territoire caractérisée par une combinaison relativement uniforme du sol et de la chronoséquence végétale.

En outre chaque unité est décrite quant aux caractéristiques morphométriques des écosystèmes aquatiques et aux matériaux géologiques de surface en bordure des étendues aquatiques.



OH4-1V1A-1-a11

REGION ECOLOGIQUE voir carte des Regions Ecologiques



SECTION REPRESENTÉE SUR LA CARTE DES TYPES ECOLOGIQUES

