

| TABLE DES MATIÈRES | |
|--|-----------|
| 1. OBJECTIF | 2 |
| 2. MISE EN CONTEXTE..... | 2 |
| 3. ATTRIBUTION DES COTES DE FIABILITÉ | 5 |
| 3.1 Cote de fiabilité sur la localisation (FL) | 6 |
| 3.2 Cote de fiabilité sur la source (FS) | 8 |
| 4. CRITÈRES DE VALIDATION DES FORAGES | 8 |
| 4.1 Validation de la localisation (VL) | 10 |
| 4.2 Validation technique (VT) | 10 |
| 4.3 Validation scientifique (VS) | 11 |
| 5. IDENTIFICATION DES DOUBLONS | 11 |
| 5.1 Sélection des doublons probables | 13 |
| 5.2 Comparaison de la stratigraphie traversée par les doublons probables | 14 |
| 5.3 Vérification et validation des doublons sélectionnés | 14 |
| 6. LA SÉLECTION DES FORAGES POUR CHAQUE LIVRABLE..... | 15 |
| 6.1 Validation des forages pour chaque livrable | 15 |
| 6.2 Choix parmi les forages doublons | 18 |
| 7. CONCLUSION | 20 |
| 8. RÉFÉRENCES..... | 20 |
| 9. MODIFICATIONS APPORTÉES | 20 |
| 9.1 Version 00 | 20 |
| 9.2 Version 01 | 20 |
| 9.3 Version 02 | 21 |
| ANNEXE 1 - TABLEAUX DES LIVRABLES AVEC FIABILITÉ ET VALIDATIONS..... | 22 |
| ANNEXE 2 - CODE VBA POUR LA COMPARAISON DE LA PROFONDEUR DU ROC..... | 23 |
| ANNEXE 3 - CODE VBA POUR LA RECHERCHE DES DOUBLONS..... | 24 |

| Version | Date | Description | Préparé par | Révisé par | Organisation |
|--|------------|---|-------------------------|--|--------------|
| <i>Rédaction initiale du protocole</i> | | | | | |
| 00.M | 2009-10-01 | Ébauche initiale du protocole | M. Lambert J. Walter | A. Rouleau R. Daigneault | UQAC |
| <i>Révision du protocole par les projets régionaux</i> | | | | | |
| 00.S | | | | | INRS |
| 00.A | | | | | UQAT |
| 00.B | | | | | UQAM |
| 00.T | | | | | UQTR |
| <i>Protocole adopté et révisions</i> | | | | | |
| 01 | (à venir) | (intégration des corrections suggérées) | M. Lambert J. Walter | R. Chesnaux A. Rouleau R. Daigneault | UQAC |
| 02 | | | | | |

1. Objectif

L'objet du présent protocole est de proposer une méthode pour l'évaluation de la fiabilité et la validation des forages dans le cadre du *Programme d'acquisition de connaissances sur les eaux souterraines du Québec* (PACES) du Ministère du Développement durable, de l'Environnement et des Parcs (MDDEP). Grâce à cette évaluation de la qualité des données, il devient possible de sélectionner les forages¹ les plus pertinents pour la création des livrables demandés dans le cadre du PACES. Le niveau de fiabilité et d'exactitude des données de forages utilisées dans un livrable sont des indicateurs de la qualité de ce livrable. Des membres du *Groupe de Recherche Interuniversitaire sur les Eaux Souterraines* (GRIES) ont élaboré ce protocole (voir tableau ci-haut) qui s'applique à tous les projets régionaux du PACES.

2. Mise en contexte

Le protocole de fiabilité et de validation des données de forage se retrouve dans le premier type de protocoles rédigés dans le cadre du PACES. En effet, quatre familles de protocoles ont été identifiées dans le cadre du PACES :

- 1 – Les protocoles portant sur l'intégration des données existantes,
- 2 – Les protocoles portant sur les méthodes de traitement des données,

¹ Le terme forage regroupe les puits d'exploitation, les piézomètres et les forages d'exploration

- 3 – Les protocoles portant sur les techniques de levés de terrain et
- 4 – Les protocoles portant sur l'édition des divers produits à élaborer.

La première famille comprend deux protocoles : a) l'archivage et la numérisation des données existantes et b) la fiabilité et la validation des forages. Pour parvenir aux livrables exigés par le PACES, le GRIES a élaboré une procédure visant à uniformiser la numérisation et la sélection finale des forages. La Figure 1 illustre ce processus et permet de situer la première famille de protocole dans les étapes définies par le GRIES, de même que les étapes spécifiques au présent protocole. La Figure 1 est brièvement décrite dans ce qui suit.

La première étape consiste à réaliser une collecte des données existantes de forage (étape 1, Figure 1). Les données de forage proviennent de différentes **sources** telles que les rapports hydrogéologiques papiers, la base de données du MTQ et la base de données du SIH. Ces données sources sont numérisées et/ou intégrées dans des fichiers *Excel* sous forme de **listes de données**, c'est-à-dire que les informations sont structurées sous forme de colonnes et de lignes. Une ligne contient donc l'information décrivant un forage en fonction des champs descriptifs structurés en colonnes. Les détails de l'étape de numérisation et de l'archivage (étape 2, Figure 1) se trouvent dans le protocole *Numérisation et archivage* rédigé par le GRIES. Déjà au moment de la numérisation dans une liste de données, les forages sont évalués selon la fiabilité de la source et la fiabilité des coordonnées permettant leur localisation (étape 3, Figure 1). La « fiabilité » porte sur la méthode d'obtention de la donnée et elle est indiquée à l'aide de **cotes de fiabilité**. Cette étape fait l'objet de la section 3 du présent protocole. Ensuite, les données sont transférées dans une base de données centrale dont la structure est commune à toutes les régions qui participent au PACES. Différents traitements sont effectués à l'aide de codes informatiques (*Visual Basic Applications*) ou de requêtes SQL (*Structured Query Language*) sur la donnée structurée en base de données. Ces codes correspondent aux étapes 4 à 6 (Figure 1) et sont présentées ci-dessous :

- Étape 4 : Traitement qui consiste à valider les informations techniques et/ou scientifiques de chaque forage. Cette étape est décrite à la section 4 du présent protocole.
- Étape 5 : Traitement visant à **identifier les doublons** de forages dans la base de données. À cette étape, un identifiant unique est attribué à chaque couple de doublons (*IdDoublon*). Les détails de ce traitement sont présentés à la section 5 du présent protocole.
- Étape 6 : Traitement crucial permettant la sélection finale des forages qui permettront l'élaboration d'un livrable donné. Cette étape est présentée à la section 6 du présent protocole. Elle repose notamment sur le **Tableau 4 de la section 6**.

La Figure 1 montre aussi les données qui seront recueillies sur le terrain au cours de la seconde phase du PACES (à l'été 2010). Ces données sont saisies à l'aide de fiches et sont directement ajoutées dans la base de données centrale. Les cotes de fiabilité sont attribuées à ces nouvelles données au moment de l'ajout dans la base de données centrale. Quelques livrables sont présentés à titre indicatif au bas de la Figure 1.

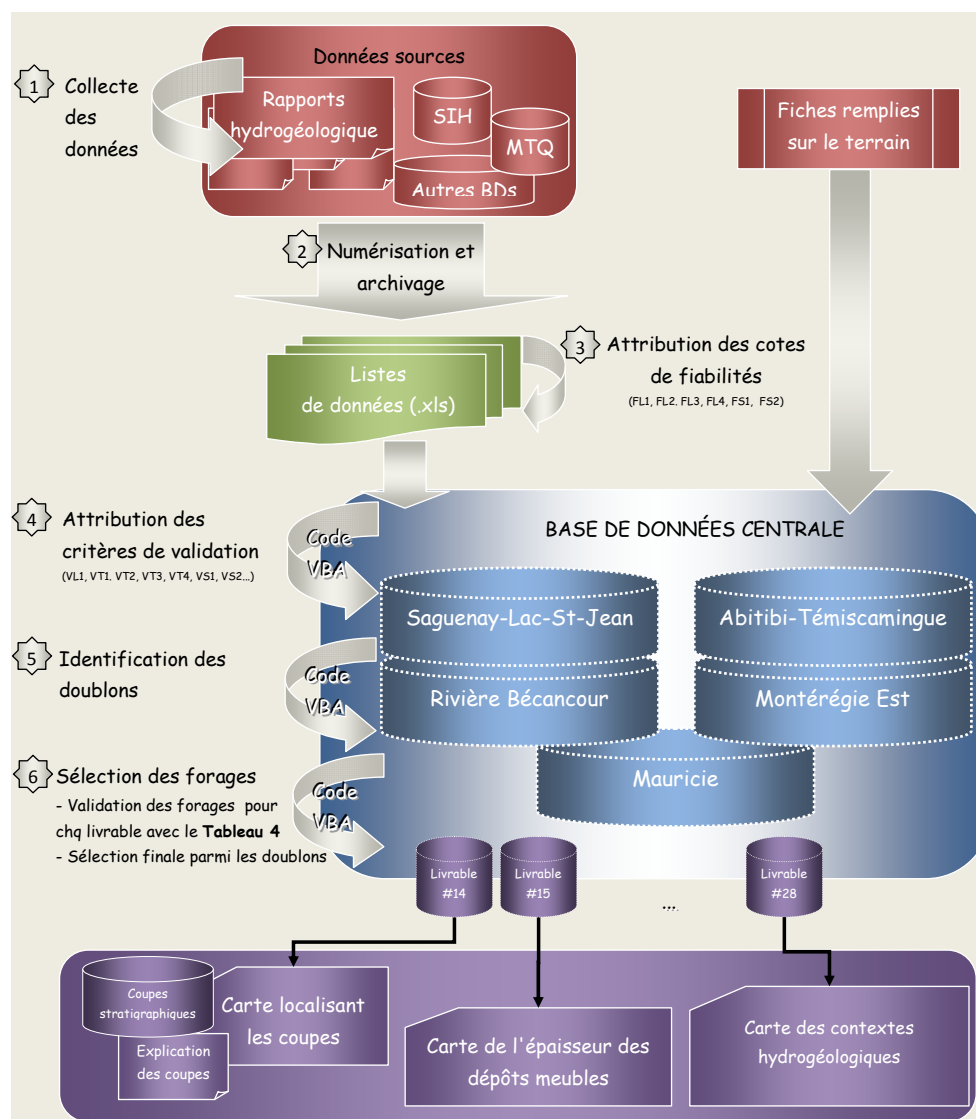


Figure 1 - Diagramme du processus de numérisation et de sélection des forages pour les livrables

3. Attribution des cotes de fiabilité

Les cotes de fiabilité ont pour objectif d'évaluer le degré d'incertitude dans la localisation, la provenance et la qualité de l'information rattachée à un forage. Ces cotes sont obtenues en considérant uniquement la méthode utilisée pour obtenir l'information. Les cotes de fiabilité ne tiennent pas compte de la valeur (magnitude) affichée pour cette information. Un champ pour chaque cote de fiabilité (cf.

Tableau 1) est ajouté dans la liste de données.

Deux catégories de cote de fiabilité ont été définies leurs noms ont été choisis pour faciliter leur identification :

- FL – Cote de Fiability sur la Localisation ;
- FS – Cote de Fiability sur la Source ;

Pour les deux cotes, une valeur élevée indique un degré d'incertitude élevé en la donnée de forage alors qu'une valeur faible indique un degré d'incertitude faible. Le

Tableau 1 présente les facteurs d'estimation et les valeurs possibles des cotes de fiabilité de la localisation et des cotes de fiabilité de la source des données. Le

Tableau 1 se retrouve aussi en format *Excel* dans le répertoire *Annexe 1 - tableaux des livrables avec fiabilité et validation*. Les sections 3.1 et 3.2 décrivent en détail ces cotes de fiabilité.

Tableau 1 - Les cotes de fiabilités

| | Cote de fiabilité | Description | Valeurs possibles |
|--------------|-------------------|---|--|
| LOCALISATION | FL1 | Source des coordonnées | 1 - Nouvelles coordonnées prises sur le terrain dans le cadre du projet PACES 2 - Les coordonnées proviennent du document source 3 - Erreur évidente dans les coordonnées sources (ex: Y= 5 332 444 au lieu de 5 232 444). L'erreur a été corrigée. |
| | FL2 | Degré d'incertitude sur les coordonnées XY | 1 - Précision entre <u>0m et 1m</u> (ex : GPS haute précision, arpentage) 2 - Précision entre <u>1m et 10m</u> (ex : GPS simple) 3 - Précision entre <u>10m et 100 m</u> (ex : extraites par géoréférencement d'une carte, extraites d'une carte topographique, localisation visuelle) 4 - Inconnu |
| | FL3-A | Degré d'incertitude sur le géoréférencement de la carte | 1 - Les coordonnées XY fournies 2 - Géoréférencement excellent 3 - Géoréférencement bon 4 - Géoréférencement acceptable 5 - Aucun géoréférencement, localisation visuelle |
| | FL3-B | Degré d'incertitude sur la localisation visuelle | 1 - Les coordonnées XY fournies ou carte géoréférencée 2 - Repères visuels fiables situés à proximité du forage 3 - Repères visuels éloignés 4 - Repères visuels peu fiables ou inexistants |
| SOURCE | FS1 | Provenance de l'information (Descriptive, stratigraphique et technique) | 1 - Données extraites de rapports techniques obtenus dans le cadre du projet 2 - Forages réalisés dans le cadre du PACES 3 - Données extraites d'une compilation de données de même source 4 - Données extraites d'une compilation de sources multiples 5 - SIH 6 - Source de données douteuses |
| | FS2 | Accessibilité des données sources (rapport originaux) | 1 - Les valeurs ont été vérifiées à l'aide des données sources 2 - Les valeurs ne peuvent être vérifiées à l'aide des données sources (source non-disponible) |

3.1 Cote de fiabilité sur la localisation (FL)

La localisation est une des caractéristiques les plus importantes du forage. Sa fiabilité (précision et exactitude) détermine l'utilisation des informations qui s'y rattachent. La première section du Tableau 1 présente le détail de cette cote.

La cote FL1 permet d'identifier les forages dont les coordonnées sources ont été modifiées. Cette cote permet d'indiquer que les coordonnées sources ont été corrigées pour obtenir une localisation plus précise ou exacte. Si les coordonnées sources sont modifiées suite à une nouvelle prise des coordonnées sur le terrain dans le cadre du projet PACES, le forage est coté 1. Les forages dont les coordonnées n'ont pas été modifiées et correspondent aux coordonnées inscrites dans le document source sont cotés 2. Lorsqu'une erreur évidente est identifiée dans les coordonnées sources au cours de la numérisation (ex: le consultant a inscrit pour la coordonnées Y: 5 332 444 au lieu de 5 232 444) et que cette erreur a été corrigée de façon fiable, le forage est coté 3. Cette opération est possible seulement lorsque des informations supplémentaires sont disponibles dans le document source décrivant le forage (ex: coordonnées d'un puits voisin, carte localisant le forage, etc.). Les erreurs flagrantes dans les coordonnées peuvent aussi être identifiées à l'aide d'un système d'information géographique.

La cote FL2 permet d'évaluer le degré d'incertitude dans les coordonnées sources lorsque ces dernières sont fournies dans le rapport. Ces coordonnées spatiales sont parfois obtenues à l'aide d'un GPS (données précises et exactes, forage coté 1), extraite d'une carte manuellement (peu précis, forage coté 3). Il arrive aussi que l'origine des coordonnées d'un forage soit inconnue, par exemple lorsque ces dernières sont le produit d'une transformation du système de référence original. Dans ce cas, le forage est coté 4.

La cote FL3 est une caractéristique de la localisation manuelle des forages, lorsque les coordonnées ne sont pas présentées avec la donnée source. Elle a été divisée en deux cotes distinctes (FL3-A et FL3-B) pour tenir compte des deux façons de faire des partenaires du GRIES : 1) localisation à l'aide du géoréférencement des cartes fournies avec les données sources (FL3-A) et 2) localisation, sans géoréférencement, à l'aide de repères visuels présentés sur la carte fournie avec la donnée source. Lors du géoréférencement d'une carte, l'erreur *RMS (root mean square)* est un bon indicateur de la précision du géoréférencement, toutefois, à lui seul le *RMS* ne suffit pas pour caractériser la qualité du géoréférencement dans son ensemble. Dans bien des cas, le nombre de points de contrôle utilisés et la transformation polynomiale choisie (de 1^{er} ordre, de 2^{ème} ordre ou de 3^{ème} ordre) doivent aussi être pris en compte. Pour simplifier l'évaluation de la qualité d'un géoréférencement, une appréciation qualitative du résultat final est réalisée en comparant la localisation des principaux repères présents sur la carte suite au géoréférencement de leur réelle localisation (ex: Base de Données Topographique du Québec (BDTQ)), tout en considérant le *RMS*, le nombre de points et l'ordre de transformation polynomiale. Ces deux cotes sont décrites ci-dessous.

La cote FL3-A indique un degré d'incertitude dans le géoréférencement d'une carte localisant des forages à l'aide d'un système d'information géographique (SIG). La carte provient habituellement d'un rapport ou d'un autre document source. Lorsque les coordonnées du forage sont fournies dans le rapport et ont été obtenues par une autre méthode, le forage est coté 1. Une cote de 2 signifie que la carte possède suffisamment de détails pour que le géoréférencement de la carte puisse être réalisé de façon précise et exacte. Une cote de 4 signifie que le géoréférencement de la carte a été plutôt difficile, mais qu'il est tout de même acceptable. Lorsque que le géoréférencement est impossible ou jugé inutile et que le forage est seulement localisé à l'aide de repères visuels, le forage est coté 5.

La cote FL3-B indique un degré d'incertitude lors d'une localisation visuelle du forage d'après certains repères présents sur la carte. La localisation visuelle du forage est utilisée lorsque la carte ne possède pas suffisamment de détails pour être géoréférencée (ex. plan à main levé du site) ou que le géoréférencement est jugé inutile (ex : un seul puits est localisé sur la carte). Une cote de 2 signifie que plusieurs repères visuels sont présents à proximité (ex. intersection de routes, hydrographie, etc) et que le forage est bien localisé. Une cote 3 signifie que des repères visuels sont présents sur la carte mais sont éloignés du site de forage. Enfin, une cote de 3 signifie que les repères visuels ne sont pas fiables ou sont inexistantes sur la carte. Une cote de valeur 1 signifie que les coordonnées du forage ont été obtenues par géoréférencement d'une carte ou que les coordonnées étaient fournies directement dans le rapport (ex: coordonnées obtenues par GPS).

N.B. : Aucune cote de fiabilité sur l'élévation du forage n'est spécifiée à cette étape-ci car la valeur de l'élévation du forage qui sera utilisée dans la création des livrables sera dérivée d'un modèle numérique d'élévation (ex: levé lidar, MNA 1:20 000, etc) et non pas celle fournie dans les données sources.

3.2 Cote de fiabilité sur la source (FS)

Le contenu scientifique des données de forage varie d'une source d'information à l'autre. Pour que l'utilisateur de la donnée puisse en faire une interprétation convenable, il convient d'évaluer la pertinence de cette information. Les cotes de fiabilité sur la source permettent d'évaluer l'accessibilité des informations attachées à un forage et l'exactitude de ces informations. Les détails de ces cotes de fiabilité sur la source sont présentés à la deuxième section du Tableau 1.

La cote FS1 indique la provenance de l'information rattachée à un forage. Les données provenant des rapports techniques sont considérées plus fiables (forage coté 1) que les données provenant de la base de données du Système d'informations hydrogéologiques du MDDEP (SIH) (forage coté 5).

La cote FS2 indique à l'utilisateur si les données sources sont accessibles et peuvent être consultées au besoin. Cette cote permet notamment de préciser la fiabilité des informations provenant de compilations. En effet, lorsque les données sources ayant fait l'objet d'une compilation sont disponibles, elles peuvent être vérifiées.

4. Critères de validation des forages

Les critères de validation ont pour objectif de permettre à l'utilisateur de la donnée de forage de juger de son exactitude quel que soit son niveau de fiabilité. Les critères de validation correspondent à des tests réalisés sur les données sources pour s'assurer de leur exactitude. Selon les valeurs obtenues pour les critères de validation, l'utilisateur peut décider d'écarter ou de conserver l'information qui est rattachée à un forage. Lorsqu'un forage est considéré valide pour la création d'un livrable, c'est qu'il a obtenu les valeurs minimales pour les critères de validation préétablis pour ce livrable (cf. Tableau 4).

Trois catégories de critères de validation ont été définies par le GRIES:

- VL – Critères de Validation sur la Localisation (VL);
- VT – Critères de Validation Technique (VT);
- VS – Critères de Validation Scientifique (VS).

Les noms de cotes ont été choisis pour faciliter leur identification. Pour les trois validations, une valeur élevée indique une donnée douteuse alors qu'une valeur faible indique une donnée exacte. Le Tableau 2 présente les facteurs d'estimation et les valeurs possibles des critères de validation. Le Tableau 2 se retrouve aussi en format *Excel* dans le répertoire *Annexe 1 - tableau x des livrables avec fiabilité et validation*. Les sections 4.1, 4.2 et 4.3 décrivent en détail ces critères de validation.

Tableau 2 - Les critères de validation

| | Critère de validation | Description | Opération | Valeurs possibles |
|--------------|-----------------------|---|--|--|
| LOCALISATION | VL1 | Cohérence de la localisation Conflit avec le réseau hydrographique | Validation de la localisation avec un SIG en vérifiant si le forage est localisé dans un polygone du réseau hydrographique. | 1 - Aucun conflit spatial identifié 2 - Conflit spatial avec le réseau hydrographique, mais moins de 10 m de la rive 3 - Conflit spatial avec le réseau hydrographique, mais plus de 10 m de la rive |
| TECHNIQUE | VT1 | Erreur sur l'épaisseur des couches stratigraphiques | Calcul de la différence entre la profondeur totale indiquée et le total de l'épaisseur des séquences stratigraphiques | 1 - Différence \leq 1 m 2 - $1 \text{ m} < \text{Différence} \leq 5 \text{ m}$ 3 - Différence $> 5 \text{ m}$ |
| | VT2 | Comparaison de la profondeur du roc | Un conflit est noté lorsque la différence entre la profondeur du roc indiquée par un forage et celle indiquée par un autre forage situé à moins de 50 mètres est supérieure à 1 mètre. | 1 - Aucun conflit 2 - Conflit avec un seul forage 3 - Conflit avec deux forages 4 - Conflit avec trois forages et plus |
| | VT3 | Comparaison des résultats d'analyses chimiques pour un puits | | 1 - Variations cohérentes 2 - Variations non-cohérentes 3 - Données insuffisantes |
| | VT4 | Erreur sur les niveaux d'eau | Compare le niveau statique avec le niveau dynamique | 1 - Niveau dynamique $<$ Niveau statique 2 - Niveau dynamique $>$ Niveau statique 3 - Données insuffisantes |
| SCIENTIFIQUE | VS1 | Balance ionique des analyses chimiques | | 1 - balance ionique $< 5\%$ 2 - $5\% < \text{balance ionique} < 10\%$ 3 - balance ionique $> 10\%$ 4 - Données insuffisantes |

4.1 Validation de la localisation (VL)

Le critère de validation sur la localisation permet de valider la position du forage. La première section du Tableau 2 en présente le détail.

Le critère VL1 permet d'identifier les forages pour lesquels la localisation est *a priori* douteuse. Par exemple, un forage destiné au pompage de l'eau souterraine pour des fins d'alimentation en eau potable se trouve rarement au milieu d'un lac ou d'un cours d'eau. Pour tenir compte de l'imprécision sur la localisation, lorsqu'un conflit spatial avec le réseau hydrographique de la BDTQ à l'échelle 1:20 000 (précision fournie environ 10m) est identifié, si le forage est situé à moins de 10 mètres de la rive, il est coté 2 alors que s'il est situé à plus de 10 mètres il est coté 3.

4.2 Validation technique (VT)

Le critère de validation technique permet de valider que les informations qui accompagnent un forage ne se contredisent pas et qu'elles ne présentent pas d'aberration lorsqu'on les compare avec l'information d'autres forages. La deuxième section du Tableau 2 présente en détail les six critères de validation sur l'aspect technique identifié.

Le critère VT1 permet de vérifier si la profondeur indiquée pour un forage correspond bien à la somme des séquences stratigraphiques traversées.

Le critère VT2 permet de valider la profondeur du roc indiquée par un forage en la comparant avec celles indiquées par les forages avoisinants. La validation VT3 est réalisée en deux étapes automatisées à l'aide d'une application VBA. Le diagramme de séquence de ce code VBA est fourni dans le répertoire de l'*Annexe 2 - Code VBA Comparaison profondeur du roc avec forages voisins* fournie avec le présent document. Les explications de l'utilisation du code VBA sont fournis dans un fichier *Lisez-moi.txt* situé dans le répertoire de l'annexe 2.

1. Une première sélection des forages est réalisée en identifiant tous les forages dans un rayon de moins de 50 mètres ayant une différence de profondeur du roc supérieure à 1 mètre.
2. Le nombre de conflits pour chaque forage est comptabilisé. Les forages ayant aucun conflit avec d'autres forages sont cotés 1, les forages ayant un conflit avec un seul forage sont cotés 2, les forages ayant un conflit avec deux forages sont cotés 3 et les forages ayant un conflit avec trois forages et plus sont cotés 4.

Le critère VT3 a été développé pour identifier les aberrations dans les résultats d'analyses hydrogéochimiques. Lorsque plusieurs résultats pour les mêmes paramètres chimiques ont été obtenus pour un même puits, leur comparaison permet d'identifier des variations naturelles ou anthropiques. On entend par variations anthropiques les erreurs dans la saisie numérique des valeurs mesurées lors de l'analyse. Cette validation est seulement possible lorsque plus d'une analyse hydrogéochimique a été réalisée pour un même forage et que le résultat des analyses est fourni dans le rapport. Un forage affichant des variations jugées cohérentes est coté 1, alors qu'un forage affichant des variations incohérentes est coté 2. Lorsque les données sont insuffisantes et ne permettent pas d'effectuer cette validation, le forage est coté 3.

Le critère VT4 permet d'identifier les aberrations de saisie de mesure de niveau d'eau, la valeur du niveau dynamique étant toujours inférieure à celle du niveau statique. Lorsque les données ne permettent pas d'effectuer cette validation, les données sont qualifiées d'insuffisantes pour ce critère de validation.

4.3 Validation scientifique (VS)

Le critère de validation scientifique est obtenu suite à un pré-traitement effectué sur la donnée. Ce critère implique généralement un calcul théorique en hydrogéologie afin de valider la donnée. La troisième section du Tableau 2 présente les facteurs d'estimation et les valeurs possibles des deux critères de validation scientifiques identifiés jusqu'à maintenant.

Le critère VS1 revient à effectuer le travail de contrôle de la qualité des résultats obtenus pour une analyse hydrogéochimique. L'erreur sur la balance ionique se calcule par le rapport de la différence des anions et des cations sur leur somme, le tout est multiplié par 100. Freeze et Cherry (1979) recommandent une limite supérieure de 5% d'erreur sur la balance ionique. Pour une erreur inférieure à 5%, l'analyse est considérée acceptable, dans le cas contraire, elle est rejetée ou remise en question. Plusieurs études augmentent parfois cette limite à 10%. Pour l'application de ce critère, il faudra s'assurer préalablement que l'analyse permet cette vérification (analyse complète).

5. Identification des doublons

L'identification des doublons est réalisée en trois étapes: (1) sélection de tous les couples de forage susceptible d'être des doublons, (2) comparaison de la stratigraphie traversée par les doublons probables et (3) validation des résultats obtenus aux étapes (1) et (2). Les deux premières étapes sont totalement automatisées. Une application VBA *lit la description et la stratigraphie des forages directement dans un fichier Excel (éventuellement ce seront les forages présents dans la base de données)*. Le diagramme de séquence de ce code VBA est fourni à Annexe 3. Les explications de l'utilisation du code VBA sont fournis dans un fichier Lisez-moi.txt situé dans le répertoire de l'annexe 3. Les sections 5.1, 5.2 et 5.3 présentent ces trois étapes.

La structure des tables contenant la description des forages (localisation et profondeur totale) et les couches stratigraphiques utilisées dans le code VBA est celle retrouvée dans le fichier Excel permettant le transfert des forages dans la base de données centrale (cf. section 2). La structure de ces deux tables est illustrée sur la Figure 2 et la Figure 3. Certains champs ont été masqués pour alléger les figures.

| | B | C | D | F | BM |
|---|-----------------------|---------------------|------------|-------------------|------------------|
| 1 | IDENTIFICATION | LOCALISATION | | | FORAGE |
| 2 | # Référence | X | Y | Projection | Prof. (m) |
| 3 | SLSJ0001 | 204470.81 | 5409027.48 | MTM 7 NAD83 (m) | 13.51 |
| 4 | SLSJ0002 | 203660.54 | 5407011.38 | MTM 7 NAD83 (m) | 47.15 |
| 5 | SLSJ0003 | 201805.20 | 5407274.80 | MTM 7 NAD83 (m) | 32.43 |
| 6 | SLSJ0004 | 201287.93 | 5405614.52 | MTM 7 NAD83 (m) | 23.42 |
| 7 | SLSJ0005 | 202873.93 | 5403795.70 | MTM 7 NAD83 (m) | 42.64 |
| 8 | SLSJ0006 | 204332.36 | 5403374.43 | MTM 7 NAD83 (m) | 33.93 |

Figure 2 - Structure de la table contenant la description des forages

| | B | C | D | E | F |
|---|-----------------------|----------------------|---------------|--------------|-----------------------------|
| 1 | IDENTIFICATION | STRATIGRAPHIE | | | |
| 2 | # Référence | Séquence | De (m) | A (m) | Code facies dominant |
| 3 | SLSJ0001 | 1 | 0.00 | 4.50 | S1 |
| 4 | SLSJ0001 | 2 | 4.50 | 12.61 | F1 |
| 5 | SLSJ0001 | 3 | 12.61 | 13.51 | G1 |
| 6 | SLSJ0001 | 4 | 13.51 | 13.51 | R |
| 7 | SLSJ0002 | 1 | 0.00 | 10.51 | S1 |
| 8 | SLSJ0002 | 2 | 10.51 | 45.95 | F1 |
| 9 | SLSJ0002 | 3 | 45.95 | 47.15 | G1 |

Figure 3 - Structure de la table contenant la stratigraphie des forages

Le

Tableau 3 contient une description pour chacun des champs retrouvés sur le Figure 2 et la Figure 3.

Tableau 3 - Description des champs de la table des attributs contenant les forages

| NOM DU CHAMP | DESCRIPTION |
|----------------------|---|
| # Référence | Identifiant unique de chaque forage. |
| X | La longitude du forage |
| Y | La latitude du forage |
| Prof. (m) | Profondeur totale du forage en mètres |
| Séquence | La séquence stratigraphique décrite pour le forage |
| De (m) | La profondeur où débute la couche stratigraphique |
| A (m) | La profondeur où se termine la couche stratigraphique |
| Code facies dominant | La lithologie dominante décrivant la couche stratigraphique |

5.1 Sélection des doublons probables

Premièrement, les couples de forage sont identifiés en se basant sur trois paramètres principaux : (a) la distance euclidienne entre les deux forages, (b) la différence de profondeur des forages et (c) le nombre de couches stratigraphiques

- a) La distance maximale entre deux forages permettant de soupçonner que ces deux forages peuvent représenter le même forage a été fixée à 200 mètres. Cette distance est arbitraire et peut facilement être modifiée au besoin.
- b) Une fois que le paramètre a) est satisfait, la différence de profondeur entre les deux forages permet de valider si le couple de forages situés à proximité l'un de l'autre peut être le même forage. La différence de profondeur maximale a été établie à 1 mètre. Cette valeur est également arbitraire et peut être modifiée au besoin.
- c) Une fois les critères a) et b) satisfaits, le troisième critère permet d'éliminer plusieurs couples de doublons et de retenir seulement les couples affichant le même nombre de couches stratigraphiques.

Deux forages situés à moins de 200 mètres l'un de l'autre, ayant une différence de profondeur inférieure à un mètre et le même nombre de couches stratigraphiques seront soupçonnés d'être les mêmes. Le code qui permet cette première sélection (cf. Annexe 3) peut évidemment être modifié pour utiliser des paramètres différents d'identification des doublons.

Tous les forages doivent se retrouver dans les deux structures des fichiers illustrées sur la Figure 2 et la Figure 3. Pour lancer la recherche des doublons, il suffit d'exécuter le code VBA de l'Annexe 3. L'application VBA demandera à la fin de la procédure de sauvegarder le fichier *Excel* contenant tous les couples de forages répondant aux paramètres a), b) et c).

Les doublons probables sélectionnés à la première étape sont enregistrés dans un fichier Excel et subiront une vérification avant d'être officiellement considérés comme doublon (cf. section 5.3).

5.2 Comparaison de la stratigraphie traversée par les doublons probables

La deuxième étape consiste à identifier parmi les couples de forages sélectionnés à la première étape (cf. section 5.1), les couples de forages qui sont les plus susceptibles d'être des doublons en comparant la stratigraphie des forages. Si la séquence stratigraphique traversée par les deux forages est exactement la même, il est fort probable que ce soit de réels doublons. Deux critères doivent être considérés dans la comparaison de la stratigraphie : a) la description des couches stratigraphiques et b) l'épaisseur de ces couches stratigraphiques:

- a) La description des couches stratigraphiques est facile à comparer puisque la description stratigraphique de tous les forages a été codifiée selon une légende unique. Si les deux forages n'affichent pas les mêmes descriptions à toutes les couches stratigraphiques, ils ne sont certainement pas des doublons.
- b) Une marge d'erreur de 0.2 mètre est acceptée lors de la comparaison de l'épaisseur des couches stratigraphiques. Cette marge d'erreur vise principalement à absorber les erreurs de conversion d'unité (ex: pieds convertis en mètres). Les couches stratigraphiques affichant la même description, mais n'ayant pas les mêmes épaisseurs pour les couches stratigraphiques, ne sont pas étiquetées comme doublons.

Les doublons sélectionnés à la deuxième étape sont enregistrés dans un fichier Excel et devront être vérifiés. Cette vérification fait l'objet de la prochaine section (cf. section 5.3)

5.3 Vérification et validation des doublons sélectionnés

Étant donné que les critères sont plus stricts à la deuxième étape (cf. section 5.2), seulement un petit échantillon de doublons pourra subir une vérification complète. Les résultats de cette vérification pourront ensuite simplement être reportés sur l'ensemble des doublons retrouvés dans le fichier Excel.

Les doublons sélectionnés à la première étape qui ne se retrouvent pas dans le fichier final de la deuxième étape devront subir une vérification complète pour ainsi éliminer toutes les possibilités que des doublons puissent s'y retrouver.

Suite à cette validation des doublons, un indicateur devra permettre d'identifier rapidement tous les doublons. Un champ *IdDoublons* serait donc peuplé à l'aide de valeurs séquentielles (ex. 1, 2, 3...) indiquant chaque paire de doublons. Ce champ unique sera nécessaire lors du choix final des forages parmi les doublons (cf. section 6.2) pour ainsi rapidement vérifier si des doublons sont sélectionnés pour un livrable donné (cf section 6).

6. La sélection des forages pour chaque livrable

La sélection des forages est réalisée en deux étapes :

- la validation des forages pour chaque livrable (section 6.1) et
- la sélection des forages parmi les doublons (section 6.2).

6.1 Validation des forages pour chaque livrable

La sélection des forages pour chaque livrable consiste à valider l'information décrivant le forage en spécifiant les valeurs des critères de validation minimales requis pour un livrable donné. Le Tableau 4 présente les valeurs minimales des critères de validation requis pour l'élaboration des livrables du projet PACES [GRIES (2009)]. Le Tableau 4 est aussi fourni en format *Excel* dans le répertoire *Annexe 2 - tableau x des livrables avec fiabilité et validation*. Ces valeurs minimales ont été choisies en tenant compte de l'échelle à laquelle les livrables finaux doivent être créés, c'est-à-dire à une échelle régionale.

La sélection finale des forages qui serviront à l'élaboration d'un livrable repose alors entièrement sur les informations du Tableau 4. Pour participer à l'élaboration d'un livrable donné, un forage doit présenter la valeur minimale requise pour l'ensemble des critères listés dans le Tableau 4, à la ligne du livrable considéré. Un champ par livrable (ex: *LIVRABLE15*) est ajouté dans la base de données contenant tous les puits et sondages. Ce nouveau champs est peuplé à l'aide d'une condition codée en VBA (*Visual Basic for Applications*) qui distingue les forages valides pour ce livrable, des forages non-valides. Le diagramme de la Figure 4 résume les principales actions de la sélection finale des forages pour chaque livrable demandé dans le cadre du PACES.

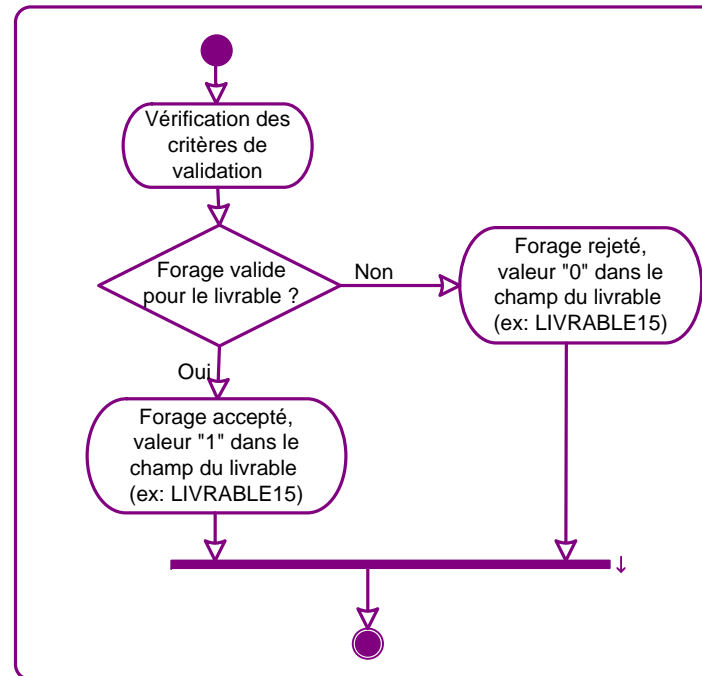


Figure 4 - Diagramme d'activité pour la sélection finale des forages pour un livrable

Le Tableau 4 contient deux acronymes :

na : signifie que la cote de fiabilité ou le critère de validation ne s'applique pas au livrable. En fait, cela signifie que les données des forages ne sont pas utilisées lors de la création du livrable.

nd : signifie que la cote de fiabilité ou le critère de validation traite d'une information qui n'est pas utile (non-déterminant) dans la création du livrable. La valeur de la cote ou du critère n'a alors aucune influence sur le choix de conserver ou de ne pas conserver les données du forage pour la création du livrable.

PROTOCOLE FIABILITÉ ET VALIDATION DES DONNÉES DE FORAGE

Version 02

Tableau 4 - Liste des livrables avec les critères de validation

| # | LIVRABLES | Critère de validation accepté | | | | | |
|----|--|-------------------------------|-----|-----|-----|-----|--------|
| | | VL1 | VT1 | VT2 | VT3 | VT4 | VS1 |
| 1 | Topographie | na | na | na | na | na | na |
| 2 | Routes, limites municipales et toponymie | na | na | na | na | na | na |
| 3 | Modèle numérique de terrain (MNT) | na | na | na | na | na | na |
| 4 | Pente du sol | na | na | na | na | na | na |
| 5 | Hydrographie | na | na | na | na | na | na |
| 6 | Limites de bassins et de sous-bassins | na | na | na | na | na | na |
| 7 | Occupation du sol | na | na | na | na | na | na |
| 8 | Couverture végétale | na | na | na | na | na | na |
| 9 | Milieux humides (zones d'intérêt écologique) | na | na | na | na | na | na |
| 10 | Affectation du territoire | na | na | na | na | na | na |
| 11 | Pédologie | na | na | na | na | na | na |
| 12 | Géologie du Quaternaire | na | na | na | na | na | na |
| 13 | Géologie du roc | na | na | na | na | na | na |
| 14 | Coupes stratigraphiques et hydrostratigraphiques | 1 | 1 | 1 | nd | nd | nd |
| | | 2 | | | | | |
| | | 3 | | | | | |
| 15 | Épaisseur des dépôts meubles | 1 | 1 | 1 | nd | nd | nd |
| | | 2 | 2 | | | | |
| 16 | Topographie du roc | 1 | 1 | 1 | nd | nd | nd |
| | | 2 | 2 | | | | |
| 17 | Contextes hydrogéologiques | 1 | 1 | 1 | nd | nd | nd |
| | | 2 | | | | | |
| 18 | Épaisseur et limites des aquifères régionaux | 1 | 1 | 1 | nd | nd | nd |
| | | 2 | 2 | | | | |
| 19 | Piézométrie des formations superficielles | 1 | nd | nd | nd | 1 | nd |
| | | 2 | | | | | |
| 20 | Piézométrie dans le roc | 1 | nd | nd | nd | 1 | nd |
| | | 2 | | | | | |
| 21 | Paramètres hydrogéologiques | 1 | 1 | 1 | nd | 1 | nd |
| | | 2 | | | | | |
| 22 | Vulnérabilité des aquifères selon la méthode DRASTIC | na | na | na | na | na | na |
| 23 | Activités potentiellement polluantes | na | na | na | na | na | na |
| 24 | Qualité (critères eau potable) | na | na | na | 1 | na | 1 2 |
| 25 | Qualité (objectifs esthétiques) | na | na | na | na | na | na |
| 26 | Utilisation de l'eau | na | na | na | na | na | na |
| 27 | Stations météo, hydrométriques et de suivi de la nappe | na | na | na | na | na | na |
| 28 | Zones de recharge préférentielles et de résurgence | 1 | nd | nd | nd | 1 | nd |
| | | 2 | | | | | |

na : ne s'applique pas

nd: non déterminant. N'a aucune influence sur le livrable

Exemple d'application du Tableau 4 :

Le livrable #15, carte de l'épaisseur des dépôts meubles, les valeurs retrouvées dans le Tableau 4 indiquent que les valeurs acceptables pour le critère de validation VL1 (*Cohérence de la localisation*) sont "1" ou "2". Un forage doit donc nécessairement posséder une valeur de zéro ("1") ou deux ("2") sous le champ VL1 pour entrer dans la création du livrable #15. Le principe demeure le même pour les autres critères de validation. Dans ce cas, le code VBA pour peupler le champ LIVRABLE15 (Livrable #15 "Carte de l'épaisseur des dépôts meubles") est le suivant :

```

IF
    VL1 <=2 AND
    VT1 <= 2 AND
    VT2 =1
THEN
    LIVRABLE15 = 1
ELSE
    LIVRABLE15 = 0
END IF
    
```

Une valeur de 1 pour le champ *LIVRABLE15* indique que les informations du forage sont valides pour être utilisées dans la construction du livrable #15, la carte de l'épaisseur des dépôts meubles. Une valeur de 0 indique que les informations du forage ne répondent pas aux exigences du livrable.

6.2 Choix parmi les forages doublons

Le diagramme de la Figure 5 résume les grandes étapes dans le choix final des forages parmi les doublons pour un livrable donné. Il est nécessaire pour la plupart des livrables d'éliminer cette duplication de données, car le poids des forages doublons devient alors deux fois plus grand que celui d'un forage sans doublon.

Parmi les forages sélectionnés à la première étape (section 6.1) les doublons sont identifiés à l'aide du champ *IdDoublons* (cf. section 5.2). Les cotes de fiabilités et/ou les critères de validation (cf. section 3 et 4) sont utilisés pour sélectionner les données du forage les plus fiables parmi les doublons. Entre les doublons, le forage qui est finalement conservé pour l'élaboration du livrable est identifié (nouvelle application *VBA*) de façon à pouvoir le sélectionner lors du traitement ultérieur (élaboration du livrable). Cette identification finale se concrétise dans un second champ (ex: *LIVRABLE15_FINAL*) où la valeur "1" sera attribuée à tous les forages valides pour le livrable, mais cette fois sans aucun doublon.

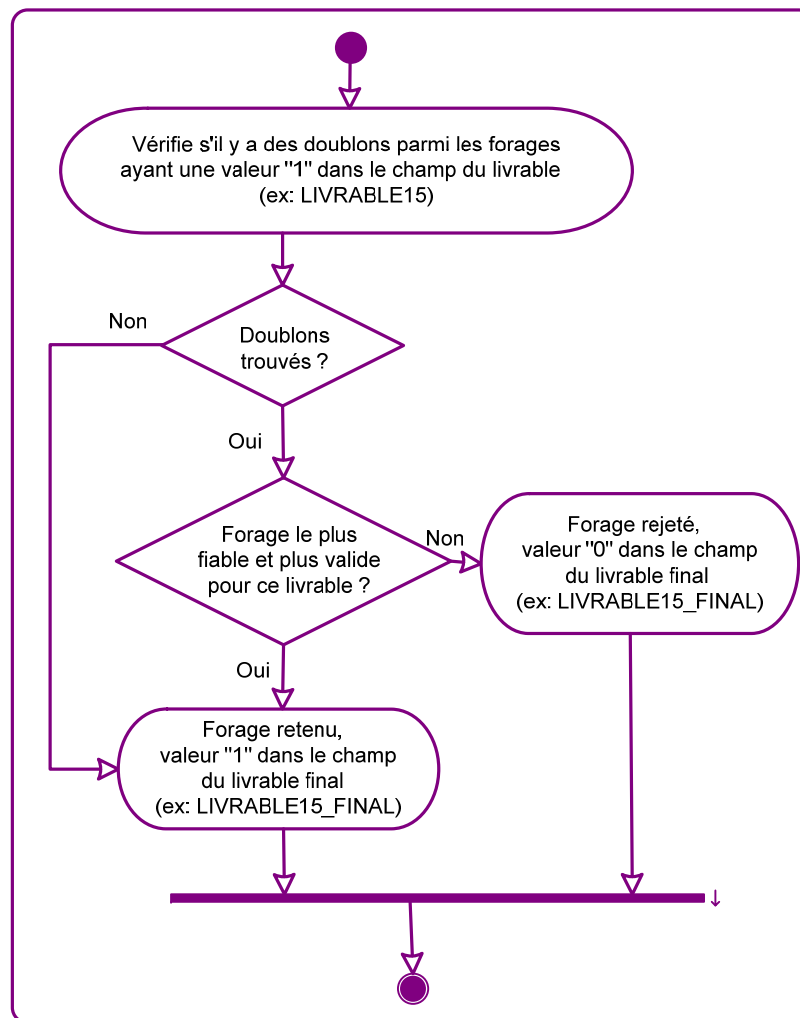


Figure 5 - Diagramme d'activité pour le choix final des forages parmi les doublons pour un livrable

7. Conclusion

L'objectif de ce protocole était d'établir une méthode opérationnelle permettant d'évaluer la fiabilité et la validité de l'information des forages qui sont compilées dans le cadre du PACES en vue de les utiliser dans l'élaboration des livrables exigés par le MDDEP. Des cotes de fiabilité et des critères de validation ont été déterminés pour permettre de qualifier chaque forage. Ces critères bien précis caractérisent d'une part la provenance de la donnée et d'autre part la donnée en soit. Le niveau de qualité des données utilisées dans la création des livrables est le principal indicateur de la précision et de l'exactitude des livrables finaux. Chaque livrable faisant appel à différentes informations rattachées à un forage, le niveau de qualité de ces informations peut varier d'un livrable à l'autre. Le **Tableau 4** (section 6.) est l'élément déterminant de ce protocole. Il présente la valeur minimale requise pour un critère de validation dans le but de discriminer les forages qui serviront à l'élaboration d'un livrable. La sélection finale des forages se fait entièrement à l'aide des valeurs présentées dans le Tableau 4. De plus, les données sur les forages provenant de sources multiples, il est essentiel de vérifier si un même forage n'est pas sélectionné à plusieurs reprises (recherche de doublons, section 6). Il est nécessaire pour la plupart des livrables d'éliminer cette duplication de données, car le poids des forages doublons devient alors deux fois plus grand que celui d'un forage sans doublon. À cette étape ultime, les cotes de fiabilité et les critères de validation servent à sélectionner le meilleur forage parmi les doublons identifiés. L'attribution des cotes de fiabilité et de validation, ainsi que l'identification des doublons et la sélection finale des forages, se font à l'aide de l'application VBA présentée tout au long de ce protocole.

8. Références

Freeze R.A., Cherry J.A. (1979) Groundwater. Englewood Cliffs, NJ: Prentice-Hall, Inc., 604p.

GRIES (2009) Cadre pour le développement des protocoles de production des livrables. Rapport de recherche, Programme d'acquisition de connaissances sur les eaux souterraines du Québec, 26p.

9. Modifications apportées

9.1 Version 00

Ébauche initiale. Les institutions qui révisent le protocole doivent modifier l'entête en indiquant le numéro de version relié à leur institution et en remplaçant « Ébauche initiale » par exemple par « Révision UQAC ».

9.2 Version 01

Les corrections importantes suggérées dans le cadre du processus de révision par les autres institutions seront décrites dans cette section.

9.3 Version 02

Si nécessaire, les corrections importantes requises pour justifier la rédaction d'un protocole révisé seront décrites dans cette section.

Annexe 1 - tableau x des livrables avec fiabilité et validations

Fichiers électroniques

Livrables avec fiabilité et validation_S20091002.xls

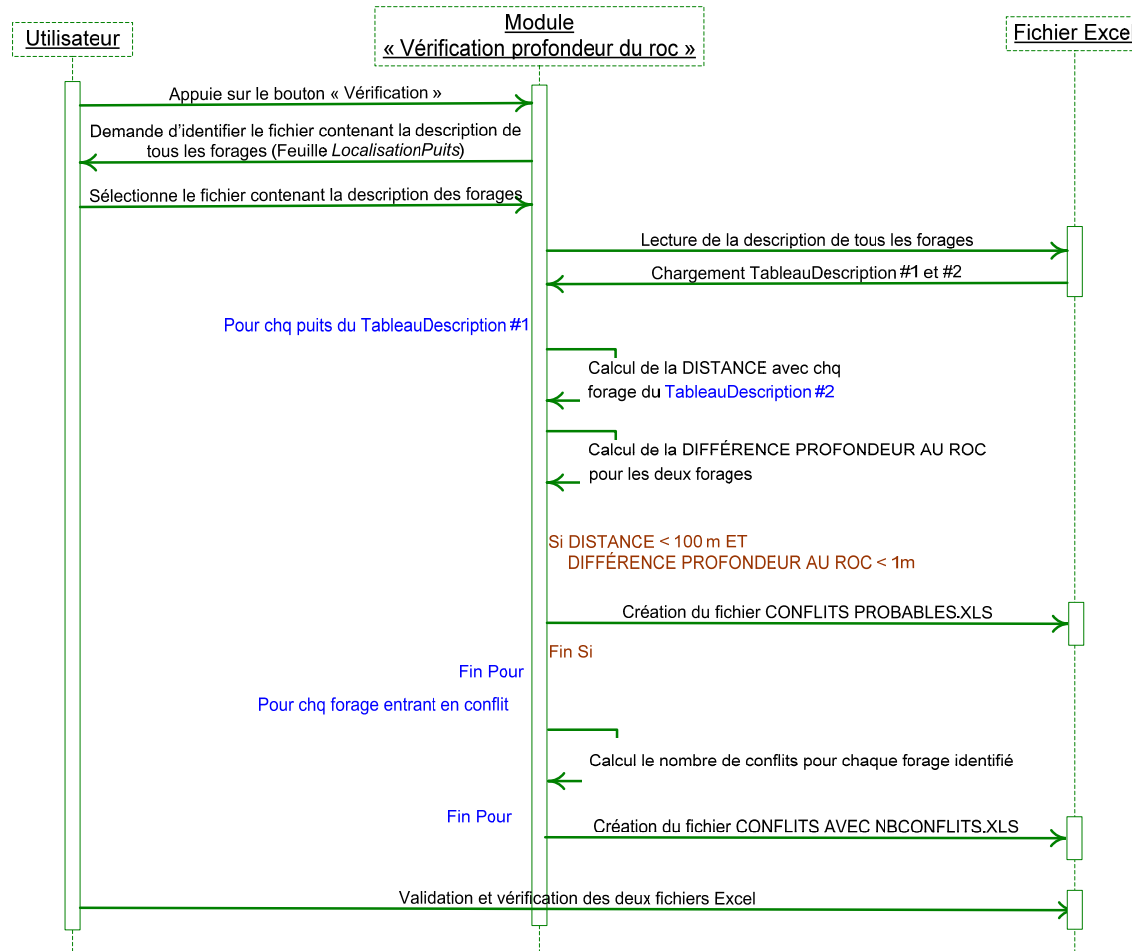
Ce fichier contient :

Tableau 1 - Les cotes de fiabilités

Tableau 2 - Les critères de validation

Tableau 4 - Liste des livrables avec les critères de validation

Annexe 2 - Code VBA pour la comparaison de la profondeur du roc



Code VBA fournit dans le fichier : *CodeVBA-ComparaisonProfondeurRoc_S20090930.txt*

Annexe 3 - Code VBA pour la recherche des doublons

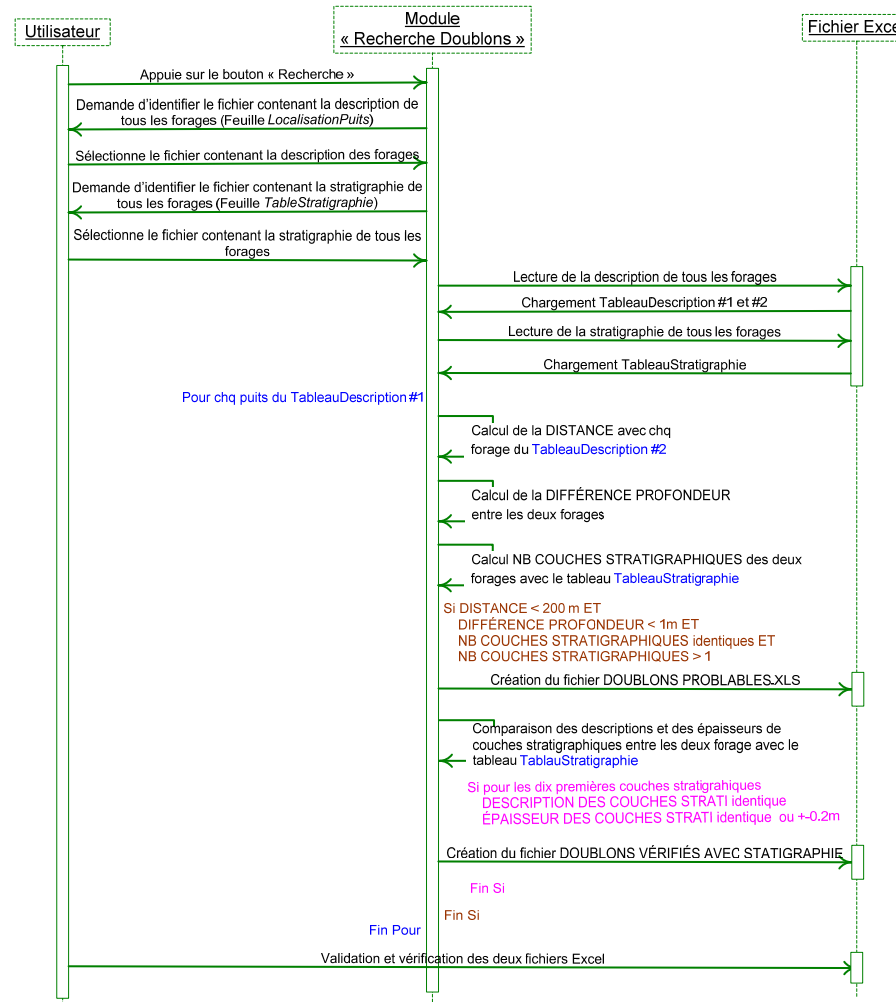


Figure 7 - Diagramme de séquence - Recherche des doublons

Code VBA fournit dans le fichier : *CodeVBA-RechercheDoublons_S20090930.txt*