

TABLE DES MATIÈRES	
1. OBJECTIF	2
2. NATURE DES DOCUMENTS	2
3. ÉTAPES DU PROTOCOLE.....	3
3.1 Étape 1 : Numérisation de premier niveau et attribution du numéro d'enregistrement (RN).....	4
3.2 Étape 2 : Saisie de données en format numérique	4
3.3 Étape 3 : Archivage	4
3.4 Étape 4 : Vérification de la numérisation de second niveau.....	5
4. RÉFÉRENCES.....	6
5. MODIFICATIONS APPORTÉES	6
5.1 Version 00	6
5.2 Version 01	6
5.3 Version 02	6

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

1. Objectif

L'objet du présent protocole est de proposer une méthode de numérisation et d'archivage des documents collectés 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). Des membres du *Groupe de recherche interuniversitaire sur les eaux souterraines* (GRIES) ont assuré la préparation de ce protocole (voir tableau ci-haut) qui s'applique à tous les projets régionaux du PACES.

2. Nature des documents

De nombreuses données concernant les eaux souterraines régionales ont été recueillies au fil des ans, à des fins diverses, par des organismes publics, gouvernementaux et municipaux, ainsi que des entreprises privées. Plusieurs de ces données ont été collectées pour des objectifs hydrogéologiques (recherche en eau, étude hydrogéologique restreinte) mais plusieurs autres portent sur des problèmes géotechniques (étude géotechnique) et environnementaux (caractérisation environnementale). Même si l'objectif original de la collecte de certaines données ne visait pas les eaux souterraines, ces données sont potentiellement utiles à la cartographie hydrogéologique, particulièrement lorsqu'elles contiennent des descriptions du sous-sol, portant soit sur la géologie, sur la géotechnique ou sur la pédologie. Ces données existantes sont donc très variées et ce sous plusieurs aspects, notamment : 1) l'objet sur lequel elles portent (géologie, débit d'eau, etc.), 2) leur qualité (précision, fiabilité, etc.), 3) leur format (tableau de valeurs, carte, etc.), 4) leur support physique (sur papier, informatique, etc.), 5) les droits de propriété, les coûts de l'obtention et la confidentialité, 6) leur année de production. De plus, ces informations utiles sont disparates, éparses, ce qui les rend difficilement accessibles et utilisables. La qualité et la fiabilité des données varient d'une source à une autre, et même d'un paramètre à un autre à l'intérieur d'une même source. Les données les plus pertinentes au projet du PACES sont de nature hydrogéologique et généralement produites par les firmes d'experts du secteur privé. L'obtention de ces études hydrogéologiques est souvent ardue, le moyen le plus simple étant de s'adresser directement au promoteur des projets qui conservent généralement une copie papier des documents.

La première phase du PACES consiste en la collecte de l'ensemble des données existantes sur le territoire d'étude. Elles sont ensuite intégrées numériquement dans une base de données à référence spatiale qui facilitera notamment la gestion de l'information collectée. Une fois compilée et collectée, la donnée source doit être rigoureusement archivée car elle servira de référence tout au long des processus de production des divers livrables exigés par le programme.

3. Étapes du protocole

Afin de faciliter la compréhension et l'application du présent protocole, quatre étapes sont proposées. Elles sont présentées sur la Figure 1 et décrites dans le texte qui suit. Dans la première étape, les documents sur support papier sont numérisés et un identifiant unique leur est attribué (cf. section 3.1). L'information utile à la production de livrable est ensuite saisie numériquement (cf. section 3.2) et l'information brute extraite du document source est archivée à des fins de contrôle (cf. section 3.3). La dernière étape en est une de contrôle et de vérification des précédentes (cf. section 3.4).

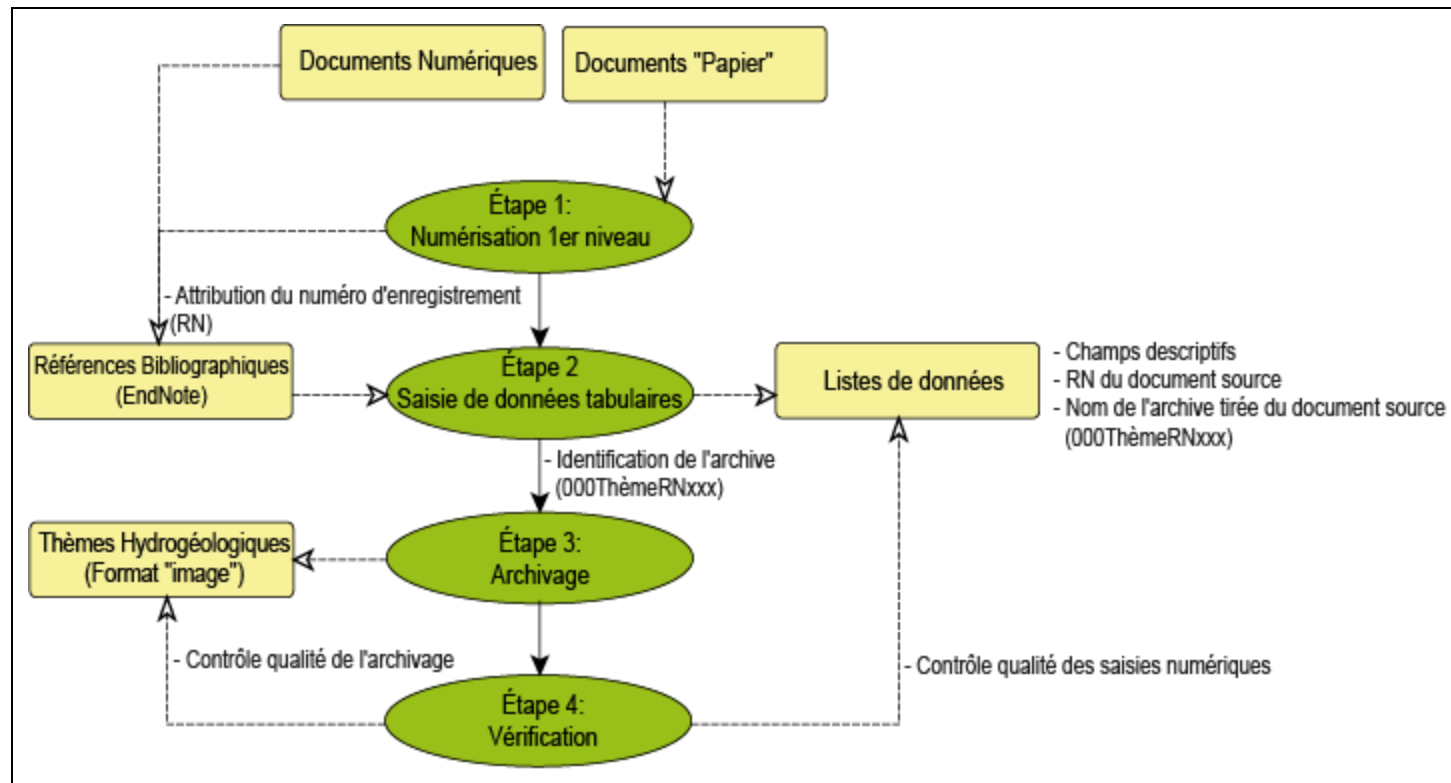


Figure 1 : Diagramme schématisant la procédure de numérisation et d'archivage des documents existants

3.1 Étape 1 : Numérisation de premier niveau et attribution du numéro d'enregistrement (RN)

Tout d'abord, les documents sources sur support papier sont numérisés à l'aide d'un balayeur optique (*scanneur*). Les documents sur support papier ainsi numérisés de même que les documents obtenus en format numérique sont ensuite stockés dans une base de données de références bibliographiques (tel que *EndNote* de Microsoft). Un numéro d'enregistrement (*Record Number*) est attribué automatiquement par le logiciel à chaque nouvelle entrée dans la banque de données bibliographique. Cet identifiant unique assure la traçabilité de la donnée source tout au long du processus d'archivage et de numérisation. Les documents constituent les archives bibliographiques du projet. Les rapports dont le contenu se rapporte au territoire d'une municipalité sont nommés avec un numéro séquentiel d'archivage, le nom de la municipalité et le numéro d'enregistrement dans la base de données bibliographiques (ex : 001RobervalRN001). Si le document porte sur une partie ou sur l'ensemble de la région, alors le nom de la municipalité sera remplacé par le titre du document.

3.2 Étape 2 : Saisie de données en format numérique

À cette étape une série d'informations jugées essentielles dans l'élaboration d'un livrable spécifique est entrée dans le chiffrier *Excel*. Les informations sont structurées sous forme de colonnes (champs) et de lignes (enregistrements). La terminologie d'*Excel* pour référer à cette structure particulière qui s'apparente à une petite base de données est une « liste de données ». Ce terme est utilisé tout au long de ce protocole et réfère à une structure telle que retrouvée à l'Annexe 1. Les champs descriptifs de la liste de données sont établis en fonction du thème hydrogéologique abordé. Un champ réservé au numéro d'enregistrement (RN) tiré de la base de données bibliographique obtenu lors de l'étape précédente doit être présent de manière à rattacher les informations au document duquel elles proviennent. Un exemple de liste de données contenant les données obtenues à partir des sondages existants est présenté à l'Annexe 1. Parmi les champs considérés dans cette liste de données il y a le nouveau nom donné au sondage, le type de sondage (puits, piézomètre, ...), le nom tiré du document source, l'année de pompage (i.e. année de forage), les coordonnées de localisation et d'élévation, la source de ces coordonnées, la profondeur totale du sondage, la profondeur du roc, le niveau statique de la nappe, la référence à la diagraphie du sondage extraite du document source, etc. À ce stade de la rédaction de ce protocole, le choix des champs retrouvés dans les listes de données est laissé à la discrétion des hydrogéologues en charge du projet. En effet, une base de données standardisée pour les fins du PACES sera bientôt disponible. Elle permet d'identifier l'ensemble des données colligées pour un thème donné. De plus, la plupart des listes de données vont inclure un ou plusieurs champs permettant de valider et de caractériser la fiabilité des informations saisies. Ces critères et ces cotes font l'objet d'un autre protocole actuellement en rédaction.

3.3 Étape 3 : Archivage

Les pages des documents numérisés à l'étape 1 (cf. section 3.1) et ayant servies à compléter la liste de données de l'étape 2 (cf. section 3.2) sont extraites des documents sources (numérisés au premier niveau) et archivées par thème (localisation, propriétés hydrauliques, diagraphie de forage, analyses hydrogéochimiques, ...). Les pages sont identifiées par un numéro séquentiel

d'extraction pour un même thème, suivi du nom du thème et du numéro d'enregistrement (RN). Les archives sont conservées dans des répertoires se rapportant aux différents thèmes abordés dans le projet. Un nouveau champ contenant le nom de l'archive est introduit dans la liste de donnée. Cette opération permet de retrouver l'information listée rapidement. Il est suggéré de procéder ainsi pour l'ensemble des listes de données construites à partir de documents numérisés au premier niveau, notamment toutes celles qui touchent aux sondages existants : localisation (000CarteRNxxx), diagraphies de forage (000DiagRNxxx), propriétés hydrauliques (000EssHydroRNxxx), résultats d'analyses chimiques (000AnalESRNxxx), aires d'alimentation (000AirAlimRNxxx), levés géophysiques (000GeophysRNxxx), ...

Cas de la localisation des sondages existants :

Plusieurs informations importantes peuvent être absentes dans le texte du document mais présentes sur des cartes et graphiques joints aux documents. Ainsi, les coordonnées de localisation d'objets ponctuels tels que des sondages par exemple, peuvent être incomplètes ou absentes dans le texte. Dans ce cas, elles peuvent être obtenues à partir des informations sur les cartes jointes au document. Lorsque c'est le cas, les cartes en format matriciel sont géoréférencées à l'aide d'un GIS et la position (i.e. localisation) des objets est saisie numériquement. La carte extraite du document source est nommée selon la convention et archivée en format matriciel dans les archives spécialisées du projet sous le thème de la localisation. Par exemple, la trentième carte du document source portant le numéro d'enregistrement (RN) 377 est nommée: 030CarteRN377 et placée dans le répertoire *ArchivesHydrogeologiques/Localisation*. Dans la liste de données sur les sondages existants (cf. Annexe 1), le champ « SourceCoor » (source des coordonnées) est peuplé avec le nouveau nom de la carte ayant permis d'obtenir les coordonnées de localisation, ce qui indique que les coordonnées ont été obtenues par géoréférencement d'une carte matriciel. Un nouveau champ numérique est alors ajouté pour indiquer le degré de confiance dans le géoréférencement de la carte (voir le *Protocole de la fiabilité et validation des forages*). Lorsque les coordonnées sont tirées du rapport, c'est le numéro d'enregistrement du rapport qui est inscrit dans le champ « SourceCoor ».

3.4 Étape 4 : Vérification de la numérisation de second niveau

Une fois les listes de données terminées, les informations saisies doivent être vérifiées minutieusement de façon à garantir l'intégrité du processus de numérisation et d'archivage des données existantes. Cette étape a pour objectif de déceler les erreurs de saisie lors de la seconde étape. Elle permet aussi de contrôler la concordance entre la donnée archivée en format matriciel (étape 3, cf. section 3.3) et celle maintenant sous forme de liste de données numériques (étape 2, cf. section 3.2). L'Annexe 2 et Annexe 3 présentent un exemple de l'étape de vérification appliquée à la liste de données sur les sondages existants, telle qu'élaborée par l'UQAC.

4. Références

5. Modifications apportées

5.1 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 ».

5.2 Version 01

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

5.3 Version 02

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

Annexe 1 – Exemples de champs spécialisés pour la liste de données des sondages existants

Microsoft Excel - BD_Description (F.MTQ.SIH-GUHS) [Lecture seule]											
	B	C	D	E	G	H	I	K	M	N	O
2	NumForage	NomForage	TypeForage	AnPompage	MunicipRap	X_MTM7	Y_MTM7	Elev	SourceCoor	ProfTotal	ProfRoc
3	F0001	"70.26"	sondage	1970	Saint-Augustin	204470,812	5409027,478		Rapports-22	13,51	13,51
4	F0002	70,27	sondage	1970	Saint-Augustin	203660,538	5407011,384		Rapports-22	47,15	47,15
5	F0003	70,28	sondage	1970	Saint-Augustin	201805,203	5407274,799		Rapports-22	32,43	32,43
6	F0004	70,29	sondage	1970	Saint-Augustin	201287,926	5405614,519		Rapports-22	23,42	
7	F0005	70,3	sondage	1970	Saint-Augustin	202873,929	5403795,703		Rapports-22	42,64	42,64
8	F0006	70,31	sondage	1970	Saint-Augustin	204332,356	5403374,433		Rapports-22	33,93	33,93
9	F0007	70,32	sondage	1970	Saint-Augustin	198429,958	5406008,177		Rapports-22	21,92	21,92
10	F0008	70,33	sondage	1970	Saint-Augustin	197560,502	5406991,072		Rapports-22	7,51	7,51
11	F0009	70,34	sondage	1970	Saint-Augustin	195363,464	5407847,765		Rapports-22	34,23	34,23
12	F0010	70,53	sondage	1970	Saint-Augustin	200041,935	5407940,053		Rapports-22	23,42	23,42
13	F0011	PZ-5	piézomètre	1999	Sainte-Hedwidge	166495,518	5373624,361	89,47	Rapports-35	7,62	
14	F0012	PZ-6	piézomètre	1999	Sainte-Hedwidge	166540,142	5373531,572	93,42	Rapports-35	16,15	
15	F0013	PZ-7	piézomètre	1999	Sainte-Hedwidge	166452,058	5373566,815	90,32	Rapports-35	9,14	

Microsoft Excel - BD_Description (F.MTQ.SIH-GUHS) [Lecture seule]																
	B	O	P	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG
2	NumForage	ProfRoc	NivStat	DiagrapRef	AnalGranul	AnalChimiq	EssaiPomp	EssaiPerma	DegConfGeo	NoProjet	ErrCoorSrc	NivDyna	DiffDynStat	VerfProf	ErrLoc	SrcCoord
3	F0001	13,51		001Diag-RN022	non	non	non	non								2
4	F0002	47,15		001Diag-RN022	non	non	non	non								2
5	F0003	32,43		001Diag-RN022	non	non	non	non								2
6	F0004			001Diag-RN022	non	non	non	non								2
7	F0005	42,64		001Diag-RN022	non	non	non	non								2
8	F0006	33,93		001Diag-RN022	non	non	non	non								2
9	F0007	21,92		001Diag-RN022	non	non	non	non								2
10	F0008	7,51		001Diag-RN022	non	non	non	non								2
11	F0009	34,23		001Diag-RN022	non	non	non	non								2
12	F0010	23,42		001Diag-RN022	non	non	non	non								2
13	F0011		-0,49	003Diag-RN035	oui	non	non	non								2
14	F0012		-2,29	004Diag-RN035	oui	non	non	non								2
15	F0013		-1,30	005Diag-RN035	oui	non	non	non								2

Annexe 2 – Étape de vérification appliquée à la liste de données sur les sondages existants

Contrôle de qualité des fichiers finaux

- 1- Vérifier les coordonnées par rapport aux sources
 - 2- Vérifier l'information présente et compléter ce qui manque à partir des rapports (élévation, profondeur, toit de la nappe...)
 - 3- Vérifier les unités de mesure, transformer tout en système métrique
Vérifier les décalages dans les coordonnées, les numéros de forages (les numéros de forages dans les fichiers «*description* »
et «*stratigraphie* » doivent correspondre)
 - 4- et «*stratigraphie* » doivent correspondre)
 - 5- transformer les valeurs de pieds en mètres
- Vérifier que le nom de la diagraphie archivée soit bien le bon et que la correspondance soit bonne avec l'info saisie (bon puits).
- 6- Vérifier que la profondeur des puits correspond à la dernière valeur de la stratigraphie
 - 7- Vérifier les diagraphies --> regroupé les pages d'un même puits en un seul fichier.
 - 8- Vérifier les doubles de puits par le nom et les coordonnées.
 - 9- Mettre les puits incomplets (stratigraphie absente) dans le dossier *ForagesIncomplets*.
 - 10- Mettre les puits non localisé (coordonnées absentes) dans le dossier *LocalisationACompléter*.

Reprendre les puits dans les dossiers du répertoire BDIncompletes (Tâches à compléter)

- 1- Relocaliser les puits non localisés
- 2- Obtenir les stratigraphies manquantes
- 3- Traiter les puits dans PuitsProblématiques

Annexe 3 – Exemple d'application de la procédure présentée à l'annexe 2

Microsoft Excel - BD-DescriptionForages-Controle_qualite												
N731												
	A	B	C	E	F	G	H	I	J	K	L	M
1	IDForages	NumForage	Nom forage	Vérifier coordonnées	Vérifier info présente et compléter	Vérifier unités mesure	Vérifier décalages	Vérifier strati-	profondeur puits v/s dernière	Vérifier les diagrammes	Vérifier les doubles	Date vérif.
720		F0776	A16/79	Coordonnées absentes, à relocaliser. Li	Pas d'élévation	x	x	x	x	x	x	05-juin-09
721		F0777	A17/79	Coordonnées absentes, à relocaliser. Li	Pas d'élévation	x	x	x	x	x	x	05-juin-09
722		F0778	A18/79	Coordonnées absentes, à relocaliser. Li	Pas d'élévation	x	x	x	x	x	x	05-juin-09
723		F0779	A19/79	x	Pas d'élévation, roc non atteint, pas de	x	x	x	x	x	x	05-juin-09
724		F0780	A20/79	x	Pas d'élévation, roc non atteint	x	x	x	x	x	x	05-juin-09
725		F0781	A21/79	x	Pas d'élévation, pas de toit	x	x	x	x	x	x	05-juin-09
726		F0782	A22/79	x	Pas d'élévation, pas de toit	x	x	x	x	x	x	05-juin-09
727		F0783	F7/69	x	Pas d'élévation, pas de toit	x	x	x	x	x	x	05-juin-09
728		F0784	F4/72	x	Pas d'élévation, pas de toit	x	x	x	x	x	x	05-juin-09
729		F0785	F2/73	x	Pas d'élévation, pas de toit	x	x	x	x	x	x	05-juin-09
730		F0786	3/72	x	Pas d'élévation, roc non atteint, pas de	x	x	x	x	x	x	05-juin-09
731		F0787	PR-2	Coordonnées absentes, à relocaliser. Li	Pas d'élévation, pas de toit	x	x	x	x	x	x	05-juin-09
732			Puits No 2/70	x	Pas traité, pas d'informations.		x				x	05-juin-09
733		F0788	No 1/69	Ce sont les mêmes puits que F0257 à f	Pas traité							05-juin-09
734		F0789	No 2/69	Ce sont les mêmes puits que F0257 à f	Pas traité							05-juin-09
735		F0790	No 3/69	Ce sont les mêmes puits que F0257 à f	Pas traité							05-juin-09
736		F0791	No 4/69	Ce sont les mêmes puits que F0257 à f	Pas traité							05-juin-09
737		F0792	No 5/69	Ce sont les mêmes puits que F0257 à f	Pas traité							05-juin-09
738		F0793	No 6/69	Ce sont les mêmes puits que F0257 à f	Pas traité							05-juin-09
739		F0794	No 7/69	En double avec F0263. Le puits F0263 était dans LocalisationACompléter mais il est complet ici au numForage F0794. Je l'ai complété et remis au F0263 et détruit de LocalisationACompléter.								08-juin-09
740		F0795	No 8/69	Mêmes coordonnées que 8A/69. Ce soit	Pas traité							08-juin-09
741		F0796	No 8A/69	Mêmes coordonnées que 8A/69	Pas d'élévation, roc non atteint, pas de	x	x	x	x	x	x	08-juin-09
742			Puits d'essai (8A/69)	x	Dans ForagesIncomplets - Aucune information, pas de stratigraphie.							08-juin-09
743		F0797	1	x	Pas d'élévation	x	x	x	x	x	x	08-juin-09
744		F0798	2	x	Pas d'élévation	x	x	x	x	x	x	08-juin-09
745		F0799	3	x	Pas d'élévation	x	x	x	x	x	x	08-juin-09
746		F0800	4	x	Pas d'élévation, pas de toit	x	x	x	x	x	x	08-juin-09
747		F0801	5	x	Pas d'élévation, pas de toit	x	x	x	x	x	x	08-juin-09
748		F0802	6	x	Pas d'élévation	x	x	x	x	x	x	08-juin-09
749		F0803	7	x	Pas d'élévation	x	x	x	x	x	x	08-juin-09
750		F0804	8	x	Pas d'élévation	x	x	x	x	x	x	08-juin-09
751		F0805	9	x	Pas d'élévation, pas de toit	x	x	x	x	x	x	08-juin-09
752		F0806	10	x	Pas d'élévation, pas de toit	x	x	x	x	x	x	08-juin-09
753		F0807	11	x	Pas d'élévation, pas de toit	x	x	x	x	x	x	08-juin-09
754		F0808	12	x	Pas d'élévation, pas de toit	x	x	x	x	x	x	08-juin-09
755		F0810	S-2	Pas de coordonnées. Dans localisation à compléter.								08-juin-09
756		F0811	PE-1	x	Pas traité.		x					08-juin-09
757		F0812	PZ-1	x	Pas traité		x					08-juin-09
758		F0813	S-1	Pas de coordonnées. Note dans le fichi	Pas traité							08-juin-09
759		F0814	S-2	Pas de coordonnées. Note dans le fichi	Pas traité							08-juin-09
760		F0815	S-3	Pas de coordonnées. Note dans le fichi	Pas traité							08-juin-09
761		F0816	S-4	Pas de coordonnées. Note dans le fichi	Pas traité							08-juin-09
762		F0817	S-5	Pas de coordonnées. Note dans le fichi	Pas traité							08-juin-09
763		F0818	S-6	Pas de coordonnées. Note dans le fichi	Pas traité							08-juin-09
764		F0819	S-7	x	Pas d'élévation, pas de toit	x	x	x	x	x	x	08-juin-09
765		F0820	S-8	x	Pas d'élévation, pas de toit	x	x	x	x	x	x	08-juin-09
766		F0821	S-9	x	Pas d'élévation, pas de toit	x	x	x	x	x	x	08-juin-09
767		F0822	S-10	x	Pas d'élévation, pas de toit	x	x	x	x	x	x	08-juin-09
768		F0823	S-11	x	Pas d'élévation, roc non atteint, pas de	x	x	x	x	x	x	08-juin-09
769		F0824	S-12	x	Pas d'élévation	x	x	x	x	x	x	08-juin-09
770		F0825	S-13	x	Pas d'élévation, roc non atteint, pas de	x	x	x	x	x	x	08-juin-09
771		F0826	S-14	x	Pas d'élévation, pas de toit	x	x	x	x	x	x	08-juin-09
772		F0827	S-15	Pas de coordonnées. Note dans le fichi	Pas traité							08-juin-09
773		F0828	S-16	Pas de coordonnées. Note dans le fichi	Pas traité							08-juin-09