

Métadonnées relatives à la qualité des données géographiques

REGISTRE FRANÇAIS

Version Projet

***Registre CNIG
octobre 2019***

Table des matières

1 Présentation du document.....	4
2 Registre.....	5
2.1 Exhaustivité.....	5
2.1.1 Taux d'exhaustivité.....	5
2.2 Cohérence logique.....	5
2.2.1 Cohérence conceptuelle.....	5
2.2.2 Cohérence au domaine de valeurs.....	6
2.2.3 Cohérence de format.....	7
2.2.4 Cohérence topologique.....	8
2.3 Précision de position.....	9
2.3.1 Précision absolue (ou externe).....	9
2.3.2 Précision relative (ou interne).....	12
2.4 Précision thématique.....	13
2.4.1 Justesse de classement.....	13
2.4.2 Justesse des attributs non quantitatifs.....	14
2.4.3 Précision des attributs quantitatifs.....	14
2.5 Qualité temporelle.....	15
2.5.1 Exactitude de la mesure temporelle.....	15
2.5.2 Cohérence temporelle.....	16
2.5.3 Validité temporelle.....	16
2.6 Généalogie.....	17
2.7 Résolution spatiale.....	18
2.8 Etendue temporelle.....	18
2.9 Dates de référence.....	20
2.9.1 Date de Création.....	20
2.9.2 Date de Révision.....	20
2.9.3 Date de Publication.....	21

Titre	Métadonnées relatives à la qualité des données géographiques
Sous-titre	Registre français
Description du document	Ce document décrit le registre français des mesures de qualité utilisable dans la partie Qualité des métadonnées Inspire.
Date	Le 17 octobre 2019
Versión	v2019-10 Première version du registre
Résumé	Ce document décrit le registre français des mesures de qualité utilisable dans la partie Qualité des métadonnées Inspire. Il fait référence aux choix opérées par le GT CNIG QuaDoGéo décrit dans les fiches méthodologiques sur la norme ISO 19157 relative à la qualité des données géographiques.
Sources	- Normes ISO 19157, ISO 19115, ISO 19119 - Guide de saisie des éléments de métadonnées de données - série de fiches méthodologiques sur la norme ISO 19157 relative à la qualité des données géographiques
Contributeurs	Ce document été produit par le groupe national GT CNIG QuaDoGéo Le groupe GT CNIG QuaDoGéo est piloté par la Commission Règles de Mise en œuvre (RMO) du CNIG, il est animé par le CEREMA et la MIG.
Rédacteurs	Arnauld Gallais – Mathieu Rajerison
Relecteurs	GT CNIG QuaDoGéo
Format	Formats disponibles du fichier : LibreOffice Writer (.odt), Adobe PDF
Diffusion	PDF sur internet
Organisme	Conseil National de l'Information Géographique (CNIG)
Langue	français
Mots-clés	Qualité, métadonnées, ISO 19157, ISO 19115
Statut du document	Validé par la Commission « Règles de mise en œuvre » du CNIG le xx/xx/2019
Licence	Le présent document est sous Licence Ouverte (Open Licence) Etalab



Suivi du document :

Juillet à septembre 2019 Première rédaction du registre, document projet

Octobre 2019

Ajout des rubriques :

- Taux d'erreur de formatage §2.2.3 (d'après M. Rajerison)

- Etendue temporelle, et dates de référence (INSPIRE)

- Exactitude de position et de radiométrie des données matricielle (ISO 19157)

1 Présentation du document

La connaissance de la qualité des données géographiques est devenue une nécessité à l'heure de la Directive européenne Inspire et de l'ouverture des données publiques.

Les ré-utilisateurs des informations géographiques s'interrogent fréquemment quant à la qualité des lots de données dont ils découvrent l'existence par les géo-catalogues et qui proposent des métadonnées en général fort peu explicatives quant à la qualité des données.

Or la méconnaissance de la qualité freine non seulement la mise à disposition des ressources existantes (le fournisseur ayant peur d'exposer une donnée imparfaite), mais également la réutilisation des données, le ré-utilisateur potentiel n'ayant pas les moyens de vérifier que la ressource ouverte identifiée dispose de la qualité suffisante pour l'usage souhaité.

Ainsi, le besoin n'est pas tant de chercher à améliorer la qualité des données, que de disposer d'une connaissance sincère de leur niveau de qualité, de les « qualifier ».

La norme ISO 19157 fournit un excellent référentiel pour la qualification des données géographiques. Cette norme reste néanmoins très technique et peu facilement abordable.

Ceci a conduit le Cerema à développer une méthode susceptible de qualifier les données conformément à la norme ISO 19157, dans le but de diminuer le coût et la complexité de l'évaluation de la qualité, tant pour le fournisseur que pour le ré-utilisateur.

Cette méthode se matérialise sous la forme d'un support pédagogique et d'une [série de fiches Cerema : Qualifier les données géographiques - Un décryptage de la norme ISO 19157](#).

Cette première étape méthodologique ayant été réalisée, le [GT CNIG QuaDoGéo](#) a ensuite été mandaté pour guider et faciliter la saisie ainsi que la transmission d'informations sur la qualité de la donnée au sein des métadonnées. Il s'agit donc d'améliorer la saisie des critères qualité dans les métadonnées ISO / Inspire en préconisant de bonnes pratiques à l'usage du producteur et de l'utilisateur.

Ceci nécessite de décrypter préalablement le contenu de la documentation ISO, identifier les champs concernés dans ISO 19115 et expliciter les métadonnées de qualité de façon à les rendre compréhensibles aux utilisateurs.

Or des critères de qualification importants tels que l'exhaustivité sont bien référencés dans ISO 19157 mais ne le sont pas dans ISO 19115.

Cependant les spécifications Inspire admettent les extensions de modèle compatibles ISO 19115.

D'autre part, les normes ISO 19157 et ISO 19115 doivent être considérées comme deux normes constituant un ensemble cohérent pour la transmission de métadonnées sur la qualité.

Autre exemple, le ~~critères références temporelles~~ particulièrement importantes ~~relatif à pour évaluer~~ l'actualité des données ~~n'est ne sont~~ pas portées par la norme ISO 19157 ~~mais~~ présentes dans ~~la norme~~ ISO 19115. Autrement dit la norme ISO 19157 n'est pas considérée comme indépendante, mais comme devant être utilisée en association avec les métadonnées ISO 19115 décrites [dans les recommandations nationales en ce qui concerne les Métadonnées INSPIRE pour les séries et ensemble de séries](#).

Dès lors, il est apparu nécessaire d'élaborer un registre français pour la transmission de métadonnées de qualité des données géographiques. Celui-ci englobe :

- le sous-ensemble de la norme ISO 19157 constitué des quelques mesures retenues dans les fiches méthodologiques (par rapport à la centaine de mesures référencées dans la norme) ;
- une extension propre aux mesures non répertoriées dans ISO 19157 mais préconisées dans les fiches pour des raisons méthodologiques et/ou pour s'adapter au profil français. Il s'agit par exemple du coefficient kappa pour la justesse de classement, et de la classe de précision de l'arrêté de 2003 pour la précision de position.
- des éléments de qualité portés par ISO 19115 donc obligatoires dans la saisie de métadonnées compatibles Inspire : la généalogie, [les références temporelles](#) et la résolution spatiale, mais absents de la norme ISO 19157.

Dans ce document, les critères et sous-critères de qualité sont présentés dans l'ordre utilisé par la norme ISO 19157 et des fiches méthodologiques sur la qualification des données géographiques.

Les éléments de qualité portés par ISO 19115 [uniquement](#) (hors ISO 19157) sont référencés en fin de document.

Ce registre ici présenté sous forme documentaire a vocation à être implémenté [sur le système de publication de registres de l'infrastructure de donnée géographiques française https://data.geocatalogue.fr/ncl/](#) en tant que registre informatique public et exploitable pour la saisie des métadonnées Inspire relative à la qualité des données géographiques.

2 Registre

Registre français pour la saisie des métadonnées Inspire relative à la qualité des données géographiques.

2.1 Exhaustivité

2.1.1 Taux d'exhaustivité

Le taux d'exhaustivité est la mesure de la qualité des données portant sur l'exhaustivité retenue par les fiches méthodologiques du GT CNIG QuaDoGéo.

Il s'agit d'une mesure qui n'est pas décrite dans la norme ISO 19157.

Ligne	Composant	Description
1	Nom	Taux d'exhaustivité
2	Alias	txEx
3	Critère / sous-critère	Exhaustivité / Taux d'exhaustivité
4	Mesure de base	taux d'erreur
5	Définition	Nombre total d'éléments en excès ou manquant dans le jeu de données par rapport au nombre total d'objets du jeu de données.
6	Description	<p>Soient :</p> <p>N_m : nombre d'objets constaté dans l'échantillon ou dans le jeu de données</p> <p>Nb_E : nombre d'objets en excédent dans l'échantillon ou dans le jeu de données</p> <p>Nb_O : nombre d'objets manquant dans l'échantillon ou dans le jeu de données</p> <p>$N_t = N_m + Nb_O - Nb_E$: nombre d'objets attendus dans l'échantillon ou dans le jeu de données</p> <p>Alors, le taux d'exhaustivité vaut :</p> $1 - \frac{(Nb_O + Nb_E)}{N_t}$
7	Paramètre	-
8	Type de valeur	Nombre réel, pourcentage
9	Structure de la valeur	-
10	Référence de la source	Fiches méthodologique « Critère d'exhaustivité » Non référencé dans ISO 19157
11	Exemple	Le jeu de données comprend 50 objets, 5 sont en excédents, 2 sont manquants. Le jeu de données devrait donc théoriquement contenir $50+2-5 = 47$ objets. Le taux d'exhaustivité est alors égal à : $1 - (7/47)$ soit : 85 %
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/txEx

2.2 Cohérence logique

2.2.1 Cohérence conceptuelle

Les anomalies que l'on peut potentiellement identifier pour ce sous-critère concernent des points qui peuvent être très variables, difficiles à quantifier et à comparer d'où la difficulté de ne ressortir qu'une seule mesure.

C'est pourquoi la fiche méthodologique du GT CNIG QuaDoGéo propose de ne pas essayer de mesurer par une valeur ou une note ce sous-critère et de se contenter d'une description littérale des anomalies constatées qui viendront compléter les mesures des trois autres sous-critères du critère de cohérence logique.

Ligne	Composant	Description
1	Nom	Non conformité aux règles du schéma conceptuel
2	Alias	NonConfConceptuelle
3	Critère / sous-critère	Cohérence logique / Cohérence conceptuelle
4	Mesure de base	comptage d'erreur, et description d'erreur
5	Définition	Comptage de tous les éléments du jeu de données qui ne se conforment pas aux règles du schéma conceptuel, et description littérale des anomalies constatées.
6	Description	Si le schéma conceptuel décrit implicitement ou explicitement des règles, elles doivent être suivies.
7	Paramètre	-
8	Type de valeur	Nombre entier, et texte libre répertoriant les anomalies constatées
9	Structure de la valeur	-
10	Référence de la source	Fiches méthodologique « Critère de cohérence logique » ISO 19157 identifiant 10
11	Exemple	-
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/nonConfConceptuelle

2.2.2 Cohérence au domaine de valeurs

La mesure de la qualité des données portant sur la cohérence conceptuelle retenue par les fiches méthodologiques du GT CNIG QuaDoGéo est le taux de conformité aux domaines de valeurs

Ligne	Composant	Description
1	Nom	Taux de conformité aux domaines de valeurs
2	Alias	txConfDomVal
3	Critère / sous-critère	Cohérence logique / Cohérence au domaine de valeurs
4	Mesure de base	taux de conformité
5	Définition	Taux de conformité aux domaines de valeurs
6	Description	Résultat de la moyenne des taux de conformités des attributs pour lesquels un domaine de valeur existe. A noter qu'en présence de données temporelles, cette moyenne intègre les résultats obtenus pour la validité temporelle des attributs de datation ou d'horodatage. Remarque : les attributs sans domaine de valeur sont par définition hors du périmètre de ce critère qualité.
7	Paramètre	-
8	Type de valeur	Nombre réel, pourcentage
9	Structure de la valeur	-
10	Référence de la source	Fiches méthodologique « Critère de cohérence logique » Moyenne des taux de conformité, cf ISO 19157 identifiant 17
11	Exemple	-
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/txConfDomVal

2.2.3 Cohérence de format

La mesure de la qualité des données portant sur la cohérence de format retenue par les fiches méthodologiques du GT CNIG QuaDoGéo est le conflit de structure physique.

Ligne	Composant	Description
1	Nom	Conflit de structure physique
2	Alias	confStructPhysiq
3	Critère / sous-critère	Cohérence logique / Cohérence de format
4	Mesure de base	Indicateur d'erreur
5	Définition	Indication que le stockage des éléments entre en conflit avec la structure physique du jeu de données
6	Description	Indication que le stockage des éléments entre en conflit avec la structure physique du jeu de données
7	Paramètre	-
8	Type de valeur	Booléen (« True » indique un conflit de structure physique)
9	Structure de la valeur	-
10	Référence de la source	Fiches méthodologique « Critère de cohérence logique » ISO 19157 identifiant 119
11	Exemple	True (le jeu de données est stocké dans un format de fichier erroné, par exemple format Shapefile au lieu de GML)
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/confStructPhysiq

Ligne	Composant	Description
<u>1</u>	<u>Nom</u>	<u>Taux d'erreur de formatage</u>
<u>2</u>	<u>Alias</u>	<u>tauxErrFormat</u>
<u>3</u>	<u>Critère / sous-critère</u>	<u>Cohérence logique / Cohérence de format</u>
<u>4</u>	<u>Mesure de base</u>	<u>Taux d'erreur</u>
<u>5</u>	<u>Définition</u>	<u>Taux d'enregistrements présentant un formatage erroné pour une ou plusieurs valeurs.</u>
<u>6</u>	<u>Description</u>	<u>Certaines valeurs doivent respecter des formats bien spécifiques telles les dates (ISO 8601), les numéros de téléphone, les références cadastrales, les adresses, les codes INSEE, etc.</u>
<u>7</u>	<u>Paramètre</u>	<u>=</u>
<u>8</u>	<u>Type de valeur</u>	<u>Nombre réel, pourcentage</u>
<u>9</u>	<u>Structure de la valeur</u>	<u>-</u>
<u>10</u>	<u>Référence de la source</u>	<u>Groupe de Travail QuaDoGéo</u>
<u>11</u>	<u>Exemple</u>	<u>0,25 : un quart des enregistrements présentent des valeurs dans un format erroné</u>
<u>12</u>	<u>Identifiant</u>	<u>https://data.geocatalogue.fr/ncl/qualiteMesures/tauxErrFormat</u>

2.2.4 Cohérence topologique

Les mesures de la qualité des données portant sur la cohérence topologique retenue par les fiches méthodologiques du GT CNIG QuaDoGéo sont :

- pour un usage « navigation sur un graphe » :
 - le taux de connexions erronées ;
- pour des usages basés sur des partitions surfaciques territoriales on retiendra les trois mesures :
 - le nombre de micro-surfaces non valides ;
 - le nombre d'erreurs d'auto-intersections non valides ;
 - le nombre d'erreurs de chevauchement.

Ligne	Composant	Description
1	Nom	Taux de connexions erronées
2	Alias	tauxConnexErr
3	Critère / sous-critère	Cohérence logique / Cohérence topologique
4	Mesure de base	taux d'erreur
5	Définition	Nombre de connexions nœud-tronçon erronées par rapport au nombre de connexions nœud-tronçon attendues supposées.
6	Description	Il existe une connexion nœud-tronçon là où différents tronçons se croisent. Ces arcs ont une relation topologique intrinsèque qui doit refléter la véritable structure du réseau. Si la connexion nœud-tronçon contredit l'univers réel, la connexion est dite erronée en ce qui concerne cette mesure de la qualité.
7	Paramètre	-
8	Type de valeur	Nombre réel, pourcentage
9	Structure de la valeur	-
10	Référence de la source	Fiches méthodologique « Critère de cohérence logique » ISO 19157 identifiant 22
11	Exemple	-
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/tauxConnexErr

Ligne	Composant	Description
1	Nom	Nombre de micro-surfaces non valides
2	Alias	nbMicroSurfErr
3	Critère / sous-critère	Cohérence logique / Cohérence topologique
4	Mesure de base	comptage d'erreur
5	Définition	Comptage de tous les éléments du jeu de données qui constituent des micro-surfaces non valides.
6	Description	Une micro-surface est une zone non voulue qui apparaît lorsque des surfaces adjacentes ne sont pas numérisées correctement. Les bords de surface adjacentes peuvent se chevaucher ou s'écarter de manière involontaire, en petites quantités, et provoquer une erreur topologique.
7	Paramètre	-
8	Type de valeur	Nombre entier
9	Structure de la valeur	-
10	Référence de la source	Fiches méthodologique « Critère de cohérence logique » ISO 19157 identifiant 25
11	Exemple	Cf. ISO 19157

12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/nbMicroSurfErr
----	-------------	---

Ligne	Composant	Description
1	Nom	Nombre d'erreurs d'auto-intersections non valides
2	Alias	nbBoucle
3	Critère / sous-critère	Cohérence logique / Cohérence topologique
4	Mesure de base	comptage d'erreur
5	Définition	Comptage de tous les éléments du jeu de données qui s'autointersectent de façon illégitime.
6	Description	Une micro-surface est une zone non voulue qui apparaît lorsque des surfaces adjacentes ne sont pas numérisées correctement. Les bords de surface adjacentes peuvent se chevaucher ou s'écarter de manière involontaire, en petites quantités, et provoquer une erreur topologique.
7	Paramètre	-
8	Type de valeur	Nombre entier
9	Structure de la valeur	-
10	Référence de la source	Fiches méthodologique « Critère de cohérence logique » ISO 19157 identifiant 26
11	Exemple	Cf. ISO 19157
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/nbBoucle

Ligne	Composant	Description
1	Nom	Nombre d'erreurs de chevauchement
2	Alias	nbChevauchErr
3	Critère / sous-critère	Cohérence logique / Cohérence topologique
4	Mesure de base	comptage d'erreur
5	Définition	Comptage de tous les éléments du jeu de données qui se chevauchent entre eux de façon illégitime.
6	Description	
7	Paramètre	-
8	Type de valeur	Nombre entier
9	Structure de la valeur	-
10	Référence de la source	Fiches méthodologique « Critère de cohérence logique » ISO 19157 identifiant 27
11	Exemple	Cf. ISO 19157
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/nbChevauchErr

2.3 Précision de position

2.3.1 Précision absolue (ou externe)

Les mesures de la qualité des données portant sur la précision absolue retenue par les fiches méthodologiques du GT CNIG QuaDoGéo sont :

- la classe de précision au sens de l'arrêté du 16 septembre 2003 ;
 - le valeur moyenne des incertitudes.
- La précision de position des données matricielles (cf. ISO 19157 § D.4.2) est décrite avec :
- la « valeur moyenne des incertitudes » (ISO 19157 § D.4.1.1 identifiant 28)
 - ou avec « l'erreur circulaire avec un seuil de signification de 95 % (ISO 19157 § D.4.1.3 identifiant 45)

Ligne	Composant	Description
1	Nom	Classe de précision au sens de l' arrêté du 16 septembre 2003
2	Alias	ClassePrecisionArrete2003
3	Critère / sous-critère	Précision de position / Précision absolue
4	Mesure de base	-
5	Définition	L'arrêté du 16 septembre 2003 porte sur les classes de précision applicables aux travaux topographiques réalisés par l'Etat, les collectivités locales et leurs établissements publics ou exécutés pour leur compte. Il permet au donneur d'ordre de spécifier la précision souhaitée, indépendamment des moyens mis en œuvre et met en place les modalités de son contrôle. Ainsi, il simplifie les relations entre donneurs d'ordre et maîtres d'œuvre tout en fournissant des outils partagés pour spécifier la précision, documenter les méthodes de contrôle et en comprendre les résultats.
6	Description	Il repose sur la définition d'une précision demandée qui se traduit par trois conditions à respecter. La première condition fait intervenir l'écart moyen en position, correspondant à la moyenne arithmétique des écarts en position. Les deux autres conditions sont des seuils limites déduits de probabilités statistiques : Le premier seuil peut être dépassé mais par un nombre limité de points. Le second seuil ne doit pas être dépassé par aucun écart en position Fiches méthodologique « Critère de précision de position » - annexe 1
7	Paramètre	-
8	Type de valeur	Nombre réel, mesure
9	Structure de la valeur	-
10	Référence de la source	Fiches méthodologique « Critère de précision de position » Non référencé dans ISO 19157
11	Exemple	Fiches méthodologique « Critère de précision de position » - annexe 1
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/ClassePrecisionArrete2003

Ligne	Composant	Description
1	Nom	Valeur moyenne des incertitudes
2	Alias	valMoyIncertitude
3	Critère / sous-critère	Précision de position / Précision absolue
4	Mesure de base	-
5	Définition	Valeur moyenne des incertitudes de position pour un ensemble de positions où les incertitudes de position sont définies comme la distance entre une position mesurée et ce que l'on considère comme étant la véritable position correspondante.
6	Description	Pour un nombre de point (N), les positions mesurées sont données sous forme de coordonnées x_{mi} , y_{mi} et z_{mi} (si trois dimensions). Un ensemble correspondant de coordonnées x_{ti} , y_{ti} et z_{ti} est considéré comme représentant les positions véritables en 3D :

		$e_i = \sqrt{(x_{mi} - x_{ti})^2 + (y_{mi} - y_{ti})^2 + (z_{mi} - z_{ti})^2}$ <p>Les incertitudes de position moyennes des positions absolues sont calculées selon la formule : $\bar{e} = \frac{1}{N} \left(\sum_{i=1}^N e_i \right)$</p>
7	Paramètre	-
8	Type de valeur	Nombre réel, mesure
9	Structure de la valeur	-
10	Référence de la source	ISO 19157 identifiant 28
11	Exemple	-
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/valMoyIncertitude

Ligne	Composant	Description
1	Nom	Erreur circulaire avec un seuil de signification de 95 %
2	Alias	errCirculaireP95
3	Critère / sous-critère	Précision de position / Précision absolue
4	Mesure de base	CE95 (seuil de signification de 95%, cf ISO 19157 Tableau G.5)
5	Définition	Rayon décrivant un cercle, dans lequel l'emplacement véritable du point se situe avec une probabilité de 95 %.
6	Description	<p>Les valeurs mesurées ne peuvent l'être qu'avec une certaine incertitude. En traitant la quantité mesurée comme variable aléatoire, cette incertitude peut être quantifiée avec des méthodes statistiques. Celle utilisées pour la définition de mesures de la qualité des données relatives à l'incertitude reposent sur certaines hypothèses :</p> <ul style="list-style-type: none"> - les incertitudes sont homogènes pour toutes les valeurs observées ; - les valeurs observées ne sont pas corrélées ; - les valeurs observées ont une distribution normale (gaussienne). <p>Les observations sont x_{mi} et y_{mi}. La zone de confiance est représentée par un cercle autour de la meilleure estimation de la valeur véritable. La probabilité que la valeur véritable se situe dans cette zone est calculée par une intégration de zone au dessus de la fonction de densité à deux dimensions de la distribution normale.</p> <p>La zone circulaire est caractérisée par son rayon R, utilisé comme mesure de l'exactitude de la variable aléatoire bidimensionnelle. (cf ISO 19157 § G.3.1 à 3)</p> $P(R, \sigma_X, \sigma_Y) = \frac{1}{2\pi\sigma_X\sigma_Y} \iint_{(x-x_t)^2 + (y-y_t)^2 = R^2} e^{-\frac{1}{2} \left(\frac{(x-x_t)^2}{\sigma_X^2} + \frac{(y-y_t)^2}{\sigma_Y^2} \right)} dx dy$
7	Paramètre	-
8	Type de valeur	Nombre réel, mesure
9	Structure de la valeur	-
10	Référence de la source	ISO 19157 identifiant 45
11	Exemple	-
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/errCirculaireP95

2.3.2 Précision relative (ou interne)

Les mesures de la qualité des données portant sur la précision relative retenue par les fiches méthodologiques du GT CNIG QuaDoGéo sont l'erreur verticale relative et l'erreur horizontale relative.

Ligne	Composant	Description
1	Nom	Erreur verticale relative
2	Alias	ErrVerticaleRelative
3	Critère / sous-critère	Précision de position / Précision relative
4	Mesure de base	-
5	Définition	Evaluation des erreurs aléatoires d'une entité du relief à une autre dans le même jeu de données ou sur la même carte par rapport à un référentiel vertical commun
6	Description	Les calculs sont faits sur toutes les combinaisons de paires de points possibles : - en calculant l'erreur verticale absolue pour chaque point - en calculant l'erreur relative pour chaque combinaison de paire - en calculant l'écart type relatif vertical - en calculant l'erreur verticale en fonction de l'intervalle de confiance avec les paramètres : P=90%, t=1,645 (coefficient de Student)
7	Paramètre	-
8	Type de valeur	Nombre réel, mesure
9	Structure de la valeur	-
10	Référence de la source	Fiches méthodologique « Critère de précision de position » Fiches méthodologique « Eléments statistiques » ISO 19157 identifiant 52
11	Exemple	-
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/ErrVerticaleRelative

Ligne	Composant	Description
1	Nom	Erreur horizontale relative
2	Alias	ErrHorizontaleRelative
3	Critère / sous-critère	Précision de position / Précision relative
4	Mesure de base	-
5	Définition	Evaluation des erreurs aléatoires d'une entité dans la position horizontale par rapport à une autre dans le même jeu de données ou sur la même carte.
6	Description	Les calculs sont faits sur toutes les combinaisons de paires de points possibles : - en calculant l'erreur horizontale absolue pour chaque point - en calculant l'erreur horizontale pour chaque combinaison de paire - en calculant l'écart type relatif horizontal - en calculant l'erreur horizontale en fonction de l'intervalle de confiance avec les paramètres : P=90%, t=1,645 (coefficient de Student)
7	Paramètre	-
8	Type de valeur	Nombre réel, mesure
9	Structure de la valeur	-
10	Référence de la source	Fiches méthodologique « Critère de précision de position »

		Fiches méthodologique «Eléments statistiques » ISO 19157 identifiant 53
11	Exemple	-
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/ErrVerticaleRelative

2.4 Précision thématique

2.4.1 Justesse de classement

Les mesures de la qualité des données portant sur la justesse de classement retenue par les fiches méthodologiques du GT CNIG QuaDoGéo sont :

- la matrice relative de classement erroné (RMCM) ;
- le coefficient kappa.

Ligne	Composant	Description
1	Nom	Matrice relative de classement erroné (RMCM)
2	Alias	RMCM
3	Critère / sous-critère	Précision thématique / Justesse de classement
4	Mesure de base	taux d'erreur et de conformité
5	Définition	Matrice indiquant le nombre d'éléments de la classe (i) classés sous la classe (j), divisé par le nombre d'éléments de la classe (i)
6	Description	La matrice relative de classement erroné (RMCM) est une matrice quadratique comportant n colonnes et n rangées. N indique le nombre de classes considérées $RMCM(i,j) = \frac{[\# \text{ éléments de la classe (i) classés sous la classe (j)}]}{[\# \text{ éléments de la classe (i)}]} \times 100\%$
7	Paramètre	Nom : n Définition : nombre de classes considérées Type de valeur : nombre entier
8	Type de valeur	Nombre réel, pourcentage
9	Structure de la valeur	Matrice (n x n)
10	Référence de la source	Fiches méthodologique « Critère de précision thématique » ISO 19157 identifiant 63
11	Exemple	Cf fiche méthodologique « Critère de précision thématique »
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/RMCM

Ligne	Composant	Description
1	Nom	Coefficient Kappa
2	Alias	kappa
3	Critère / sous-critère	Précision thématique / Justesse de classement
4	Mesure de base	taux d'erreur et de conformité
5	Définition	Matrice indiquant le nombre d'éléments de la classe (i) classés sous la classe (j), divisé par le nombre d'éléments de la classe (i)
6	Description	Avec les éléments de la matrice relative de classement erroné RMCM(i,j) donnés comme mesure de la qualité des données au tableau ci-dessus « Matrice relative de classement erroné », le coefficient Kappa peut être calculé comme :

		$\kappa = \frac{N \cdot \sum_{i=1}^r \text{MCM}(i,i) - \sum_{i=1}^r \left(\sum_{j=1}^r \text{MCM}(i,j) \cdot \sum_{j=1}^r \text{MCM}(j,i) \right)}{N^2 - \sum_{i=1}^r \left(\sum_{j=1}^r \text{MCM}(i,j) \cdot \sum_{j=1}^r \text{MCM}(j,i) \right)}$
		Où N est le nombre d'éléments classés
7	Paramètre	Nom : r Définition : nombre de classes considérées Type de valeur : nombre entier
8	Type de valeur	Nombre réel, compris entre -1 et 1
9	Structure de la valeur	-
10	Référence de la source	Fiches méthodologique « Critère de précision thématique » ISO 19157 identifiant 64
11	Exemple	Cf fiche méthodologique « Critère de précision thématique »
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/kappa

2.4.2 Justesse des attributs non quantitatifs

La mesure de la qualité des données portant sur la justesse des attributs non quantitatifs retenue par les fiches méthodologiques du GT CNIG QuaDoGéo est le taux de valeurs d'attributs correctes.

Ligne	Composant	Description
1	Nom	Taux de valeurs d'attributs correctes
2	Alias	tauxValAttOk
3	Critère / sous-critère	Précision thématique / Justesse des attributs non quantitatifs
4	Mesure de base	taux de conformité
5	Définition	Nombre de valeurs d'attributs correctes par rapport au nombre total de valeurs d'attribut.
6	Description	-
7	Paramètre	-
8	Type de valeur	Nombre réel, pourcentage
9	Structure de la valeur	-
10	Référence de la source	Fiches méthodologique « Critère de précision thématique » ISO 19157 identifiant 66
11	Exemple	-
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/tauxValAttOk

2.4.3 Précision des attributs quantitatifs

La mesure de la qualité des données portant sur la précision des attributs quantitatifs retenue par les fiches méthodologiques du GT CNIG QuaDoGéo est la mesure d'incertitude de la valeur d'attribut avec un seuil de signification de 95 %.

[Les valeurs de bandes des données matricielles sont également décrites avec cette mesure \(cf. ISO 19157 § D.4.2\).](#)

Ligne	Composant	Description
1	Nom	Incertitude de la valeur d'attribut avec un seuil de signification de 95 %
2	Alias	incertitudeValAttS95
3	Critère / sous-critère	Précision thématique / Précision des attributs quantitatifs
4	Mesure de base	LE95 ou LE95(r), en fonction de la procédure d'évaluation
5	Définition	Demi-longueur de l'intervalle défini par une limite supérieure et inférieure, dans laquelle la valeur véritable de l'attribut quantitatif se situe avec une probabilité de 95 %.
6	Description	Cf. ISO 19157 - Annexe G « Mesure de base de la qualité des données » §3.2 « Variable aléatoire unidimensionnelle » Cf. fiche méthodologique « Eléments statistiques » §7 « Incertitude »
7	Paramètre	-
8	Type de valeur	Mesure
9	Structure de la valeur	-
10	Référence de la source	Fiches méthodologique « Eléments statistiques » Fiches méthodologique « Critère de précision thématique » ISO 19157 identifiant 71
11	Exemple	Fiches méthodologique « Eléments statistiques » §9.2
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/incertitudeValAttS95

2.5 Qualité temporelle

2.5.1 Exactitude de la mesure temporelle

La mesure de la qualité des données portant sur l'exactitude de la mesure temporelle retenue par les fiches méthodologiques du GT CNIG QuaDoGéo est la mesure de l'exactitude de la mesure temporelle avec un seuil de signification de 95 %.

Ligne	Composant	Description
1	Nom	Exactitude de la mesure temporelle avec un seuil de signification de 95 %
2	Alias	exactMesTempoS95
3	Critère / sous-critère	Qualité temporelle / Exactitude de la mesure temporelle
4	Mesure de base	LE95 ou LE95(r), en fonction de la procédure d'évaluation
5	Définition	Demi-longueur de l'intervalle défini par une limite supérieure et inférieure, dans laquelle la valeur véritable de l'instance temps se situe avec une probabilité de 95 %.
6	Description	Cf. ISO 19157 - Annexe G « Mesure de base de la qualité des données » §3.2 « Variable aléatoire unidimensionnelle » Cf. fiche méthodologique « Eléments statistiques » §7 « Incertitude »
7	Paramètre	-
8	Type de valeur	Mesure
9	Structure de la valeur	-
10	Référence de la source	Fiches méthodologique « Eléments statistiques » Fiches méthodologique « Critère de qualité temporelle » ISO 19157 identifiant 57
11	Exemple	Fiches méthodologique « Critère de qualité temporelle » §2.1
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/exactMesTempoS95

2.5.2 Cohérence temporelle

La mesure de la qualité des données portant sur la cohérence temporelle retenue par les fiches méthodologiques du GT CNIG QuaDoGéo est la mesure du taux de conformité chronologique.

Il s'agit d'une mesure qui n'est pas décrite dans la norme ISO 19157.

Ligne	Composant	Description
1	Nom	Taux de conformité chronologique
2	Alias	tauxConfChrono
3	Critère / sous-critère	Qualité temporelle / cohérence temporelle
4	Mesure de base	taux de conformité
5	Définition	nombre d'objets ordonnés de manière correcte par rapport aux autres éléments (ou dont les attributs sont ordonnés de manière incorrecte) par rapport au nombre total d'objets du jeu de données.
6	Description	-
7	Paramètre	-
8	Type de valeur	Nombre réel, pourcentage
9	Structure de la valeur	-
10	Référence de la source	Fiches méthodologique « Critère de qualité temporelle » Non référencé dans ISO 19157 (le plus proche est cohérence temporelle – indicateur d'erreur identifiant 159)
11	Exemple	Fiches méthodologique « Critère de qualité temporelle » §4
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/tauxConfChrono

2.5.3 Validité temporelle

La mesure de la qualité des données portant sur la validité temporelle retenue par les fiches méthodologiques du GT CNIG QuaDoGéo est la mesure du taux de conformité au domaine de valeur.

Ligne	Composant	Description
1	Nom	Taux de conformité aux domaines de valeurs
2	Alias	txConfDomVal
3	Critère / sous-critère	Qualité temporelle / Validité temporelle, issu de : Cohérence logique / Cohérence au domaine de valeurs
4	Mesure de base	taux de conformité
5	Définition	Taux de conformité aux domaines de valeurs
6	Description	Résultat de la moyenne des taux de conformités des attributs pour lesquels un domaine de valeur existe. Cette moyenne intègre les résultats obtenus pour la validité temporelle des attributs de datation ou d'horodatage.
7	Paramètre	-
8	Type de valeur	Nombre réel, pourcentage
9	Structure de la valeur	-
10	Référence de la source	Fiches méthodologique « Critère de cohérence logique » Moyenne des taux de conformité, cf ISO 19157 identifiant 17

11	Exemple	-
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/txConfDomVal

2.6 Généalogie

Aux mesures de la qualité des données retenues par les fiches méthodologiques du GT CNIG QuaDoGéo par rapport à la norme ISO 19157 s'ajoutent deux mesures définies par le paragraphe « Qualité » de la norme ISO 19115 : la généalogie, et la résolution spatiale :

Ligne	Composant	Description
1	Nom	Généalogie
2	Alias	genealogie
3	Critère / sous-critère	- « Généalogie » dans ISO 19115 - (non présent dans ISO 19 157)
4	Mesure de base	-
5	Définition	La généalogie fait état de l'historique du traitement et/ou de la qualité générale de la série de données géographiques. Le cas échéant, elle peut inclure une information indiquant si la série de données a été validée ou soumise à un contrôle de qualité, s'il s'agit de la version officielle (dans le cas où il existe plusieurs versions) et si elle a une valeur légale.
6	Description	La généalogie de la ressource décrit l'historique d'un jeu de données et, s'il est connu, le cycle de vie de celui-ci, depuis l'acquisition et la saisie de l'information jusqu'à sa compilation avec d'autres jeux et les variantes de sa forme actuelle. Il s'agit d'apporter une description littérale et concise soit de l'histoire du jeu de données, soit des moyens, procédures ou traitements informatiques mis en œuvre au moment de l'acquisition du jeu de données. Par exemple, la généalogie peut consigner l'échelle de saisie si cette information est importante pour l'utilisation du jeu de données. La date ou la version de la donnée source est également un élément utile. Un lien vers une ressource documentaire externe (ce peut être un PDF) décrivant par exemple un processus de transformation, pourra être inséré.
7	Paramètre	-
8	Type de valeur	Texte libre
9	Structure de la valeur	-
10	Référence de la source	Guide de saisie des éléments de métadonnées de données - § VII.1
11	Exemples	1/ Mise au standard COVADIS PPR des fichiers MapInfo de la Préfecture de police de Paris utilisés pour la fabrication des annexes du PPRI de Paris. 2/ La méthode utilisée, la méthode hydrogéomorphologique, donne une description de la plaine alluviale fonctionnelle des cours d'eau, façonnée par leurs crues successives, en délimitant les différentes structures morphodynamiques qui la composent. Elle permet de cartographier l'enveloppe maximale de la zone inondable, et fournit également des éléments relatifs à la dynamique de la crue. Cette méthode est complétée par une approche historique (repères de crue, enveloppes de zones inondées, données d'archives ou issues de témoignages). Ces atlas sont constitués d'une cartographie sous SIG à l'échelle du 1/25 000ème voire du 1/10 000ème pour les secteurs urbanisés, ainsi que d'un rapport de présentation du bassin versant qui apporte des commentaires de la cartographie pour chaque secteur homogène identifié. 3/ Réseau numérisé sur un EDR (élément de reproduction) au 1 : 25 000 de 1992.
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/genealogie
	ISO/TS 19139 path	dataQualityInfo/*/lineage/*/statement

	Remarque	L'élément scope>level doit être fixé à « dataset » dans le cas d'une série, « series » pour un ensemble de séries. De plus, il ne doit pas y avoir de restrictions sur l'étendue spatiale.
--	----------	--

2.7 Résolution spatiale

Ligne	Composant	Description
1	Nom	Résolution spatiale
2	Alias	resolutionSpatiale
3	Critère / sous-critère	- « Résolution spatiale » dans ISO 19115 - (non présent dans ISO 19157)
4	Mesure de base	-
5	Définition	La résolution spatiale décrit le niveau de détail de la ressource.
6	Description	La résolution spatiale est exprimée comme un ensemble de valeurs de distance de résolution allant de zéro à plusieurs valeurs ou exprimée en échelles équivalentes : - Une échelle équivalente : - est exprimée sous la forme d'une valeur entière correspondant au dénominateur de l'échelle. - est utilisée en général pour les cartes ou les produits dérivés de cartes. - Une distance de résolution : - est exprimée sous la forme d'une valeur associée à une unité de longueur. - est utilisée en général pour des données maillées et des produits dérivés d'imagerie.
7	Paramètre	-
8	Type de valeur	Texte libre
9	Structure de la valeur	-
10	Référence de la source	Guide de saisie des éléments de métadonnées de données - § VII.2
11	Exemples	pour une échelle équivalente : - dénominateur : 5000 - pour une distance de résolution : - valeur : 2 - unité de mesure : m (Contre-exemples : 1/50000, 50 000, 50000°, 50000eme, 2 mètres)
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/resolutionSpatiale
	ISO/TS 19139 path	identificationInfo[1]/*/spatialResolution/*/equivalentScale/*/denominator

2.8 Etendue temporelle

Aux mesures de la qualité des données retenues par les fiches méthodologiques du GT CNIG QuaDoGéo par rapport à la norme ISO 19157 s'ajoutent deux mesures définies par le paragraphe « référence temporelle » de la norme ISO 19115 :

- l'« **Etendue temporelle** » définit la période de temps couverte par le contenu de la ressource, ce qui est un indicateur de l'actualité de la ressource :

- Les « **Dates de référence** » : sont représentatives de la vie de la ressource « **Date de création** », « **Date de dernière révision** », « **Date de publication** ».

La directive européenne INSPIRE impose la fourniture d'au moins une référence temporelle qui peut être une étendue

temporelle ou une date de référence (de création, de dernière révision, ou de publication).

Le système de référence par défaut est le calendrier grégorien, et les dates sont exprimées conformément à la norme ISO 8601 (c'est-à-dire sous la forme : **aaaa-mm-jj**)

Les heures sont exprimées sous la forme aaaa-mm-jjThh:mi:ss+hh:mi, où les quatre derniers chiffres correspondent au décalage avec l'heure UTC. Par exemple : 2019-10-17T14:42:00+02:00 (c'est-à-dire : 17 octobre 2019 quatorze heures et quarante-deux minutes, 0 secondes, zone UTC+2)

Ligne	Composant	Description
1	Nom	Etendue temporelle
2	Alias	etendueTemporelle
3	Critère / sous-critère	- « Etendue temporelle » dans ISO 19115 - (non présent dans ISO 19 157)
4	Mesure de base	Date, intervalle de date, mélange de date et d'intervalles
5	Définition	L'étendue temporelle définit la période de temps couverte par le contenu de la ressource.
6	Description	Cette période peut être exprimée par une date déterminée ; un intervalle de dates exprimé par la date de début et la date de fin de l'intervalle ; ou un mélange de dates et d'intervalles.
7	Paramètre	-
8	Type de valeur	Date Le système de référence par défaut est le calendrier grégorien et les dates sont exprimées conformément à la norme ISO 8601, sous la forme : aaaa-mm-jj.
9	Structure de la valeur	-
10	Référence de la source	Guide de saisie des éléments de métadonnées de données - § VI.1
11	Exemples	Période de validité d'un Plan Local d'Urbanisme. date de début : 2017-08-24 - date de fin : 2019-08-24 La BD Topo est actualisée en continu (tous les jours), dans ce cas le champ « Etendue temporelle » prendra la forme : "1997 à maintenant". L'élément "maintenant" est une valeur définie par la norme ISO 19108 avec le code "now".
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/etendueTemporelle
	ISO/TS 19139 path	identificationInfo[1]/*/extent/*/temporalElement/*/extent

2.9 Dates de référence

2.9.1 Date de Création

Ligne	Composant	Description
1	Nom	Date de création
2	Alias	dateCreation
3	Critère / sous-critère	- « Dates de référence / Date de création » dans ISO 19115 - (non présent dans ISO 19 157)
4	Mesure de base	Date
5	Définition	Date de création de la ressource.
6	Description	Date de création de la ressource. Il peut y avoir plus d'une date de publication. En absence de « Date de révision », la « Date de création » permet d'apprécier l'actualité de la ressource
7	Paramètre	-
8	Type de valeur	Date Le système de référence par défaut est le calendrier grégorien et les dates sont exprimées conformément à la norme ISO 8601, sous la forme : aaaa-mm-jj.
9	Structure de la valeur	-
10	Référence de la source	Guide de saisie des éléments de métadonnées de données - § VI.2.2
11	Exemples	Date de création d'un Plan Local d'Urbanisme : 2017-08-24
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/dateCreation
	ISO/TS 19139 path	identificationInfo[1]/*citation/*/date[./*dateType/*/text()='creation']/*date

2.9.2 Date de Révision

Ligne	Composant	Description
1	Nom	Date de révision
2	Alias	dateRevision
3	Critère / sous-critère	- « Dates de référence / Date de Révision » dans ISO 19115 - (non présent dans ISO 19 157)
4	Mesure de base	Date
5	Définition	Date de dernière révision de la ressource.
6	Description	Date de dernière révision de la ressource. Il ne doit pas y avoir plus d'une date de dernière révision, et elle n'a de sens que si elle se distingue de la date de création. La « Date de révision » permet d'apprécier l'actualité de la ressource
7	Paramètre	-
8	Type de valeur	Date Le système de référence par défaut est le calendrier grégorien et les dates sont exprimées conformément à la norme ISO 8601, sous la forme : aaaa-mm-jj.
9	Structure de la valeur	-
10	Référence de la source	Guide de saisie des éléments de métadonnées de données - § VI.2.3

11	Exemples	Date de révision d'un Plan Local d'Urbanisme : 2019-08-24
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/dateRevision
	ISO/TS 19139 path	identificationInfo[1]/*[citation]*/date[.]*/dateType/*[text()='revision']*/date

2.9.3 Date de Publication

Ligne	Composant	Description
1	Nom	Date de publication
2	Alias	datePublication
3	Critère / sous-critère	- « Dates de référence / Date de Publication » dans ISO 19115 - (non présent dans ISO 19 157)
4	Mesure de base	Date
5	Définition	Date de publication de la ressource.
6	Description	Date de publication de la ressource. Il peut y avoir plus d'une date de publication.
7	Paramètre	-
8	Type de valeur	Date Le système de référence par défaut est le calendrier grégorien et les dates sont exprimées conformément à la norme ISO 8601, sous la forme : aaaa-mm-jj.
9	Structure de la valeur	-
10	Référence de la source	Guide de saisie des éléments de métadonnées de données - § VI.2.1
11	Exemples	Dates de publication d'un PLU sur le Géoportail de l'urbanisme : 2017-09-15 et 2019-10-17
12	Identifiant	https://data.geocatalogue.fr/ncl/qualiteMesures/datePublication
	ISO/TS 19139 path	identificationInfo[1]/*[citation]*/date[.]*/dateType/*[text()='publication']*/date