

# ÉVALUATION DE LA QUALITÉ DES GÉODONNÉES ET OPERATIONALISATION DE LA DIRECTIVE INSPIRE : RETOUR D'EXPERIENCE

## GEODATA QUALITY ASSESSMENT AND OPERATIONALISATION OF THE INSPIRE DIRECTIVE: FEEDBACK

Gilles-Antoine NYS, Charline DUBOIS, Céline GOFFIN, Pierre HALLOT,  
Jean-Paul KASPRZYK, Mathieu TREFFER, Roland BILLEN

### Résumé

Au même titre que le Référentiel à Grande Echelle (RGE) français, le Service Public de Wallonie (SPW) vise à mettre en place son propre Géoréférentiel. Ce référentiel est destiné à répondre à l'établissement d'une infrastructure d'information géographique dans la Communauté européenne selon la directive INSPIRE. Il consiste notamment à assurer la disponibilité des premières géodonnées de référence dont la qualité est garantie et en cohérence avec les besoins métiers des utilisateurs. Dans cette optique, l'Unité de Géomatique de l'Université de Liège, en collaboration avec le Comité Technique du SPW du Géoréférentiel et la Direction de la Géométrie, propose une méthodologie de contrôle de la qualité pour trois géodonnées pressenties du Géoréférentiel : les axes des voiries, les emprises de bâtiments et les points d'adresse. Ce projet a permis de revenir sur la définition d'une base théorique pour la validation de la qualité des géodonnées. Cet article en présente les grandes lignes et principales conclusions.

### Mots-clés

géodonnée, qualité, normes, INSPIRE

### Abstract

*In the same way as the French Référentiel à Grande Echelle (RGE), the Walloon Public Service aims to set up its own Georepository. This repository is intended to respond to the establishment of an Infrastructure for Spatial Information in the European Community under the INSPIRE directive. It consists among other things in ensuring the availability of the first reference geodata whose quality is guaranteed and in coherence with the business needs of the users for the Walloon territory. With this in mind, the Geomatics Unit of the University of Liège, in collaboration with the Technical Committee of the Public Service of Wallonia for the Georepository and the Geometry Department, proposes a quality control methodology for three prospective geodata of the Georepository: buildings, road axes and addresses. One of the outcomes of this project was the definition of a theoretical basis for the validation of geodata quality. This article presents its main findings and conclusions.*

### Keywords

*geodata, quality, standards, INSPIRE*

## INTRODUCTION

La directive européenne INSPIRE fournit le cadre juridique pour la réutilisation des données spatiales et oblige les entités publiques à partager leurs données et métadonnées qui entrent dans le champ d'application de la directive (Directive INSPIRE, 2007). Tout ce qui peut poser problème, tel qu'une incompatibilité, une impossibilité de comparer ou des faibles niveaux de précision, géométriques comme attributaires, peut compromettre l'interopérabilité. Cependant, la directive ne rentre pas dans le détail et ne propose pas, par exemple, de seuils pour la validation de la qualité demandée. Les recommandations proposées ne se traduisent

donc pas toujours sur un plan méthodologique ou pratique. Il est donc de la responsabilité des organismes géographiques nationaux de fournir la documentation nécessaire pour la définition de la qualité des données fournies et sa validation. La balance doit être trouvée entre le cadre législatif imposé par la directive et les aspects techniques. La directive propose cependant une liste non exhaustive des critères de qualité mais limite leur définition en une dizaine de mots. Ces définitions, très succinctes, laissent encore trop de place aux interprétations et par conséquent, de mécompréhensions, d'impossibilités de comparaison entre les informations fournies par des organismes différents, etc.

Le Décret Wallon du 22 décembre 2010 relatif à l'infrastructure d'information géographique wallonne (InfraSIG) officialise la mise en place d'un Géoréférentiel wallon (InfraSIG, 2010). Le Plan opérationnel géomatique pour la Wallonie 2017-2019 (POGW) approuvé par le Gouvernement wallon le 16 février 2017 dresse les engagements et intentions wallonnes en géomatique en 22 actions (POGW, 2017). L'action 4 du POGW vise à définir le contenu du Géoréférentiel wallon et à assurer la disponibilité des premières géodonnées de référence à la qualité garantie, en cohérence avec les besoins métiers des utilisateurs. Le projet informatique de cartographie continue (PICC) a été choisi comme base pour enrichir celui-ci et de fait tirer parti de l'expérience acquise depuis l'achèvement de la couverture de la Wallonie en décembre 2014.

En étroite concertation avec le Comité Technique (CT-géoref) du SPW et la Direction de la Géométrie, le projet QualiGeo propose d'analyser le flux de travaux du processus d'analyse de qualité existant, afin de le rendre conforme aux normes en vigueur, de le consolider et de l'améliorer le cas échéant. Le projet vise non seulement à améliorer le processus du contrôle de qualité mais également à proposer des adaptations du processus d'acquisition pour les géodonnées ne répondant pas aux règles et standards internationaux. Dans une phase ultime, des recommandations sont discutées afin de boucler sur la mise en place du nouveau Géoréférentiel et ainsi d'améliorer la qualité des données de façon continue.

L'objectif de cet article est de présenter le contexte général du projet QualiGeo, ses principaux résultats et son intégration dans une démarche proactive d'assurance de la qualité engagée par les services publics wallons. Il s'agit concrètement de proposer une méthode de transposition des éléments de qualité repris dans la directive INSPIRE dans un contexte opérationnel. Ce document vise de plus à combler un vide laissé dans la documentation des instituts géographiques internationaux quant à la définition de leurs critères, seuils et méthodes de qualité. Il se décompose en trois parties successives :

La première partie contextualise le projet et s'intéresse au cadre méthodologique mis en place pour mesurer la qualité de données considérées comme prioritaires : bâtiments, routes et adresses.

La définition d'un glossaire en français reprend les termes principaux usuels en matière de qualité de l'information spatiale. Un état de l'art sur la validation de la qualité de couches vectorielles de géodonnées vient compléter ce glossaire en étudiant les critères définis par les instituts géographiques nationaux frontaliers et internationaux. Nous discutons, de plus, les diverses considérations qui ont mené à la détermination de six critères de qualité et de leurs treize sous-critères. Ces critères, issus notamment de la proposition faite dans la directive INSPIRE, sont à la fois quantitatifs et qualitatifs. Les grandes lignes de la méthodologie sont présentées sur un plan théorique.

Une fois, le cadre méthodologique et les critères définis, la seconde partie de l'article présente leur application et des exemples de résultats d'une campagne de validation. À titre d'exemple, la distinction est faite entre les territoires considérés comme urbains ou ruraux. Différentes méthodes d'échantillonnage sont appliquées pour les campagnes de contrôle, sur base de ce découpage et des critères de qualité définis dans le point précédent. La mise en place du cadre méthodologique de QualiGeo s'inscrit dans une boucle récursive d'amélioration continue des processus de production engagée par la Direction de la Géométrie avec le second Plan Stratégique Géomatique pour la Wallonie en 2017. La troisième partie de ce papier s'intéresse aux différentes étapes constitutives de cette boucle « *plan-do-check-act* » (PDCA) (*préparer; développer; contrôler et agir*) et revient sur l'intérêt de cette méthode de gestion de la qualité.

Avant de conclure, nous reprenons rapidement quelques remarques pratiques destinées à enrichir la méthode et son application concrète dans les processus de production et de contrôle. C'est l'occasion de faire la publicité d'une documentation accessible et normée pour l'information spatiale en ligne avec les prescriptions de la directive INSPIRE.

## I. CONTEXTE

Les possibilités croissantes d'utilisation de l'information géographique imposent de reconsidérer le problème de leur qualité. Il peut effectivement en découler non seulement des fautes directes mais aussi indirectes. Cette qualité impacte la fiabilité que l'on peut espérer des analyses spatiales et décisions prises au moyen de leurs résultats (Guptill et

Morisson, 1995). Que cela soit pour le professionnel ou le grand public, il est de plus en plus important de pouvoir mesurer cette qualité mais aussi de la maîtriser, de l'améliorer et dans un dernier temps de la garantir (Delavar et Devillers, 2010).

Conformément au décret InfraSIG, le futur Géo-référentiel wallon est destiné à comptabiliser un minimum de douze couches de géodonnées en modes vectoriel et raster. Dans un premier temps, il a été choisi de tester complètement trois couches vectorielles jugées prioritaires : les bâtiments, les axes de voirie et les points d'adresse.

La première étape dans un travail de cette envergure, avant toute étude ou validation de la qualité, consiste en la détermination d'un glossaire partagé et des définitions des termes employés. Ce ne sont pas moins de 140 termes qui ont été définis et collationnés dans un glossaire partagé. Cette liste a permis de lever une ambiguïté importante quant aux mots « précision / exactitude / justesse / fidélité ». Cet ensemble de termes porte à confusion encore bien trop souvent en matière de qualité géométrique. L'occasion est prise ici de revenir sur deux de ces définitions :

- Précision : Étroitesse de l'accord entre une mesure ou une estimation et l'espérance de cette mesure ou de cette estimation (parfois appelé « fidélité »). La précision se mesure au moyen de l'écart-type ou des estimateurs statistiques associés (David et Fasquel, 1997).
- Exactitude : Étroitesse de l'accord entre une mesure ou une estimation d'une grandeur et la valeur nominale de cette grandeur (parfois appelé « justesse »). L'exactitude se mesure au moyen du biais ou écart moyen en position (David et Fasquel, 1997).
- À noter qu'une mesure à la fois juste et fidèle est parfois appelée exacte (on ne parlera alors pas de précision / exactitude pour éviter tout malentendu).
- La confusion est souvent apportée lors de traduction vers le français depuis l'anglais. Alors que le terme « *precision* » existe et signifie en français « précision », le terme « *accuracy* » se traduit généralement par « précision » mais sous-tend en réalité la même signification que le terme exactitude en français.
- Sur un plan théorique, le projet a permis de répondre à trois grandes questions afin de mettre en place la méthodologie et de proposer les

premiers résultats:

- quels sont les critères qui permettent de définir la qualité de couches de données vectorielles ?
- quels sont les indicateurs relatifs à ces critères qui permettent de traduire de façon objective la qualité ?
- quels sont les seuils à respecter pour la validation d'une géodonnée en regard des critères définis et de leurs indicateurs respectifs ?

La section suivante présente les réponses à ces trois questions.

## II. CRITÈRES DE QUALITÉ

Suite à l'état de l'art dressé sur les pratiques des grandes institutions géographiques internationales et la directive INSPIRE, six critères sont retenus : la cohérence logique, l'exhaustivité, la qualité du positionnement, la qualité sémantique, la qualité temporelle et la généalogie.

La cohérence logique est subdivisée en quatre sous-critères. La cohérence du format concerne les concordances par rapport au format ou à l'encodage de l'information. La cohérence conceptuelle se rapporte au modèle conceptuel de données et traduit les définitions d'objets mais aussi les relations qui les lient. La cohérence du domaine (ou contrainte du domaine) correspond à l'utilisation de valeurs d'attributs prévues dans une plage ou une liste de valeurs définies dans le modèle conceptuel de données. La cohérence topologique correspond à la transcription des relations topologiques entre les objets d'une même couche. Les interactions potentielles entre les différentes couches font partie d'une seconde étude, on considère ici que les couches sont indépendantes.

L'exhaustivité, ou complétude, reprend les sous-critères d'excédent et d'omission. Une erreur par excédent correspond à un objet présent dans les données alors qu'il n'existe pas (ou plus) sur le terrain nominal. A contrario, un objet absent des données est une erreur par omission. L'exhaustivité ne doit pas être confondue avec l'actualité qui est définie plus loin.

Le positionnement se différencie selon qu'il soit absolu ou relatif. Tout écart entre les coordonnées d'un objet dans les données numériques et ces coordonnées dans le terrain nominal est une erreur

de positionnement absolu. Les erreurs de positionnement relatif se rapportent au non-respect des règles édictées dans les spécifications du produit de la mesure quant à l'orientation, à la situation des objets et à leur position les uns par rapport aux autres.

La qualité sémantique est décomposée en deux sous-critères : classement des objets et classement d'évaluation des attributs. Tout objet inclus dans une classe ne lui correspondant pas sur un plan sémantique est une erreur de classement pour l'objet. C'est principalement la nature même de l'objet qui est problématique dans ce cas. Au niveau de détail inférieur dans la définition de l'objet, un attribut qualitatif peut lui aussi être erroné (classement des attributs), tout comme un attribut quantitatif peut ne pas correspondre à la valeur mesurée sur le terrain nominal, ou tout du moins à son ordre de grandeur (évaluation des attributs). Une valeur aberrante est considérée comme mal évaluée.

La qualité temporelle correspond globalement à la notion d'actualité. Tout objet possédant une valeur d'attribut de datation antérieure d'un nombre d'années défini par rapport à une date fixée par le gestionnaire est considéré comme un défaut d'actualité. Ce critère ne permet cependant pas de faire état de l'objet sur le terrain nominal au moment de sa lecture dans la base de données. On saura simplement que l'objet est cohérent à l'instant  $t$  selon les règles du moment.

Enfin, pour la généalogie, l'absence d'information concernant la source de la création d'un objet, son lien avec l'objet dont il découle ou la raison de sa modification correspond à un défaut de généalogie. Ce dernier critère correspond à l'historique des données. En parallèle à la notion d'identité et à celle de la traçabilité d'un objet, c'est certainement le critère le plus complexe à mettre en place et à valider.

Dans un second temps, l'ensemble de ces critères et sous-critères sont retranscrits par des indicateurs de façon à permettre d'appréhender sur une base objective la qualité d'une couche de données. Vis-à-vis de l'exhaustivité et de la précision absolue, par exemple, il sera précisé à quel pourcentage le jeu de données est exhaustif (en excédent et en omission) et quelle est la précision du positionnement de ces objets en centimètres par rapport à un système de coordonnées de référence. De plus, l'ordre des critères n'est pas anodin. Dans le cas où un critère

ne validerait pas une donnée, il n'est pas intéressant de mobiliser des ressources pour qualifier les autres critères. À titre d'exemple, il ne serait pas utile d'étudier un objet qui existerait dans la base de données mais n'existe plus dans le monde réel (erreur par excédent) et ce, même si le format, la cohérence logique, etc. sont valides. Ainsi, la complétude est un critère à prendre en considération en amont d'autres critères.

### III. MÉTHODOLOGIE ET CONTRÔLE

Plusieurs points de la méthodologie nécessitent une explication. Cette section reprend les quelques éléments théoriques nécessaires à la bonne compréhension de la démarche proposée.

En premier lieu, l'exactitude du positionnement de l'emprise d'un bâtiment peut être déterminée à partir de trois primitives géométriques : des points, des lignes et des surfaces. Du fait de cette nature variable, différentes mesures quantitatives sont nécessaires :

- Points : l'écart moyen en position, c'est-à-dire la moyenne arithmétique des écarts en position relevés sur les points des objets géographiques, d'un échantillon de points, comparé à un levé de contrôle plus précis (issus de l'Arrêté Ministériel français du 16 septembre 2003 portant sur les classes de précision applicables aux catégories de travaux topographiques (JORF, 2003).
- Lignes : distance d'Hausdorff, écart maximal entre deux lignes, qui mesure la distance entre deux ensembles de points réciproquement (Hangouët, 1995 ; David et Fasquel, 1997 ; Vauglin, 1999). Il s'agit d'une mesure de similarité entre deux formes.
- Surfaces : indice surfacique qui fournit une indication de la distance entre deux surfaces en tenant compte de leur surface commune et des polygones parasites (Bel Hadj Ali et Vauglin, 2000 ; Vauglin, 1999).
- Les mesures sont appliquées en cascade, de la primitive la plus complexe à la plus simple (du polygone au point), et sont associées à des seuils intermédiaires de rejet afin d'invalider progressivement les entités erronées. L'idée est ici la même que lors de l'ordonnement des critères pour l'optimisation de l'utilisation des ressources. Cela permet également de vérifier l'appariement des entités à contrôler avec leurs homologues de référence.

Selon le critère considéré, il peut être intéressant de limiter les vérifications à un petit nombre d'éléments représentatifs. Une méthode d'échantillonnage basée sur la norme ISO19157:2013 a donc été mise en place pour chaque couche (International Organization for Standardization, 2013). Cet échantillonnage s'effectue à plusieurs niveaux d'échelle : (i) une grille globale de découpage de la Wallonie ; (ii) échantillonnage sur les bâtiments compris dans les mailles sélectionnées au point (i) ; (iii) échantillonnage semi-aléatoire sur les axes de voirie et les points-adresse dans les mailles sélectionnées au point (i).

Théoriquement, un échantillonnage est toujours possible mais il peut être intéressant qu'il ne concerne pas tous les critères. Prenons l'exemple du format de données, ces erreurs sont facilement intégrées dans des processus globaux informatisés tels que des processus ETL (*Extract-Transform-Load*) et ne demandent pas le même niveau d'investissement. On peut imaginer que l'automatisation de ces vérifications sur l'entièreté du jeu de données se fasse à une fréquence définie (hebdomadaire, mensuelle, annuelle, etc.).

Une fois le cadre méthodologique défini, la phase de mise en pratique a pour but, d'une part, de vérifier la faisabilité des processus proposés mais aussi, d'autre part, de proposer du matériel de communication pour la publicité des résultats.

Un modèle conceptuel de données est mis en place lors de la création des bases de données. Ce modèle est ensuite traduit dans un modèle logique qui formalise les relations entre les différents objets pour enfin être traduit dans un modèle physique, qui structure directement la base de données. Cependant, les besoins et attentes des utilisateurs changent. Il subvient à ces changements que la structuration des données évolue régulièrement elle aussi dans le temps. Des mises à jour peuvent alors être nécessaires pour gérer une nouvelle information, retirer une information obsolète, etc. Des modifications des modèles conceptuels, logiques et donc physiques sont nécessaires. Ce point est discuté dans les recommandations et perspectives du projet.

Il est donc à prévoir que les seuils de validation à atteindre varient eux-aussi. De plus, ces variations peuvent s'expliquer, non seulement en fonction de

l'évolution des besoins des utilisateurs, mais aussi des nouvelles techniques d'acquisition, du recouplement avec d'autres sources d'information, etc. Il peut donc être important de revoir les seuils tels que définis par les grands instituts géographiques internationaux, et de les mettre en regard de ce qu'il est demandé dans le cadre de la mise en place du futur Géoréférentiel. Les seuils minima à fixer suite à cet état de l'art pour la validation des jeux de données sont présentés dans le Tableau 1 en fonction des critères et sous-critères définis ci-dessus.

Ce tableau reprend les seuils théoriques à atteindre lors de la production des différentes couches de géodonnées. C'est donc une limite théorique minimum. Néanmoins, ces limites théoriques peuvent ne pas être parfaitement atteintes lors de la phase d'acquisition / production (à noter que l'on peut très bien les dépasser mais que leur coût n'est pas justifié en regard des besoins des utilisateurs mis en évidence dans la phase de conceptualisation).

Concrètement, l'étude n'avait pas pour but de définir le meilleur seuil à atteindre ou le plus difficile. Le principal résultat du projet était de revenir sur la détermination de la balance entre la qualité et la consommation de ressources pour atteindre cette qualité.

Dans une démarche de cette envergure, il était important de trouver un moyen de communication efficace pour synthétiser les conclusions de l'étude et les comparer avec les seuils fixés, mais aussi de se identifier face aux démarches des voisins. L'idée était d'obtenir en un coup d'œil une représentation globale de la qualité d'une couche de géodonnées, tout en permettant la comparaison avec d'autres couches ou d'autres sources d'information pour les mêmes couches.

Un diagramme de Kiviat, dit aussi en radar ou en étoile, est une méthode graphique efficace permettant d'afficher des données multivariées sous la forme d'un diagramme bidimensionnel de trois variables quantitatives, ou plus, représentées sur des axes partant du même point. Utilisé efficacement, un diagramme en radar permet de présenter les valeurs minimales et maximales, les valeurs à atteindre et les valeurs actuelles de tous les résultats et coûts, d'une manière accessible

Critères principaux & sous-critères (acquisition)	Seuils internationaux			
	Unité de base	Bâtiments	Axes de voiries	Points d'adresses
<b>Exhaustivité</b>				
Excédent	%	1,0	0,5	2,0
Omission	%	1,0	0,5	2,0
<b>Cohérence logique</b>				
Cohérence conceptuelle (objet/attribut)	%	0,1	0,1	0,1
Cohérence de domaine	%	0,1	0,1	0,1
Cohérence de format (objet/attribut)	%	0,1	0,1	0,1
Cohérence topologique	booléen	vrai	vrai	vrai
<b>Positionnement</b>				
Positionnement absolu (écarts aux sommets)	cm	10	30	N/A
Positionnement relatif (écarts aux sommets)	cm	20	30	N/A
<b>Qualité sémantique</b>				
Classement des objets	%	2,0	1,0	2,0
Classement des attributs	%	2,0	2,0	2,0
<b>Qualité temporelle</b>				
Qualité temporelle (actualité) (MAJ cyclique)	an	2,0	2,0	0,5
<b>Généalogie</b>				
Généalogie (jeu de données)	booléen	vrai	vrai	vrai
Généalogie (objets)	booléen	vrai	vrai	vrai

**Tableau 1.** Résumé des seuils de qualité déterminés suite à l'état de l'art

à toutes les parties prenantes. Il faut pour cela que les axes soient normés – de façon linéaire ou logarithmique – et correspondent à une réalité de coûts. L'outil peut être dès lors être utilisé pour mettre en place une stratégie d'investissement pour l'amélioration de la qualité. Il offre une vision synoptique et efficace des points faibles et points forts de la qualité d'une couche dans son ensemble, et permet de les compenser ou de repenser les financements et l'écriture de cahiers des charges en vue d'améliorations (voir Figure 1). Les variations dans le tracé des coûts et des résultats nécessiteront de futures interventions dans les groupes de discussion adéquats.

Indépendamment de la nature du critère de qualité à déterminer, un résultat notable de l'état de l'art a permis de mettre en avant les outils et méthodes préférés par les instituts nationaux géographiques. La détermination de ces outils et la balance entre leur niveau de « performance » et le coût qu'ils représentent ont permis de proposer une solution

directement assimilable dans la continuité de l'expérience des équipes du SPW. Il est certain qu'une visite sur site avec une station totale permettrait d'obtenir un niveau de qualité plus que nécessaire mais elle nécessiterait a priori un investissement en temps et en argent trop important. Des solutions telles que l'acquisition depuis un véhicule mobile offrent une solution efficace pour un coup réduit, mais la garantie de la qualité du levé de contrôle est moindre.

Mais la réalité d'un institut n'est pas celle d'un autre. Les moyens des uns ne peuvent donc pas être déterminants pour les autres, disposant d'une expérience, de compétences et d'un personnel qui leurs sont propres. Une part importante de la présente étude a donc consisté à déterminer, dans cette réalité, qu'elle était la stratégie la plus à même d'être mise en place. La volonté première était de proposer une validation de la qualité de leurs données par les services publics en interne, avant de penser à externaliser celle-ci.

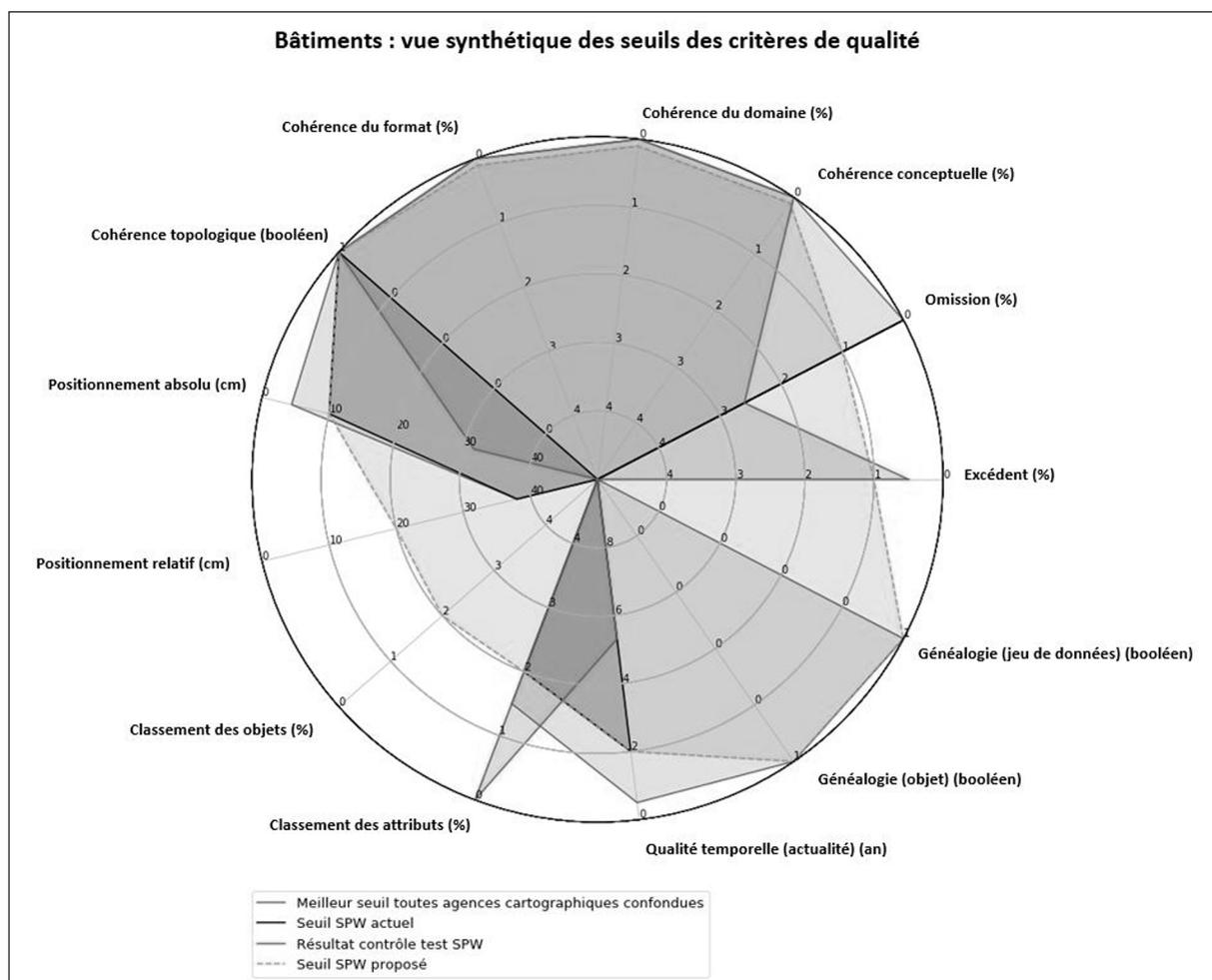


Figure 1. Exemple d'un diagramme en radar pour la couche «bâtiments» selon les critères définis

Un second point important est à relever dans la continuité de l'Arrêté Ministériel français du 16 septembre 2003 pour la validation du positionnement planimétrique (JORF, 2003): les spécifications renseignées dans l'étude sont des spécifications de résultats, et non plus des spécifications de moyens. Précédemment, il était d'usage de renseigner, dans le cahier des charges d'un marché public que cela soit pour les levés ou pour leur validation, la méthodologie de levé afin de cadenciser l'exécution dudit marché en espérant en garantir la qualité. Ceci n'est plus possible actuellement. Le nombre et la variété des méthodes d'acquisition étant toujours plus nombreux, l'obligation de résultats permet de faire fi des limitations qu'engendrait l'imposition de tel ou tel instrument ou procédé. De plus, c'est un confort non négligeable et une possibilité d'innovation certaine pour les différents prestataires. Il est alors demandé au prestataire de garantir avec certitude le niveau de précision de la méthode choisie. À titre d'exemple, on parle dans l'arrêté d'un « levé de contrôle ayant une précision au moins

deux fois supérieure à la donnée à lever » (JORF, 2003). Une méthode normée d'échantillonnage permet de garantir la confiance sur l'ensemble du jeu de données en se limitant à un petit nombre d'éléments vérifiés sur terrain. Ceci permet de limiter les coûts en ressources des vérifications, mais aussi aux prestataires de mieux intégrer ces processus dans leur pratique.

Le Tableau 2 reprend, en regard de chaque couche vectorielle et de chaque critère considérés, la méthode ou l'outil le plus adéquat pour la détermination de sa qualité selon l'état de l'art. Un filtre a cependant été appliqué pour déjà correspondre aux ressources internes des services publics wallons.

Concernant l'étude de la qualité des trois couches suivantes (le relief, les limites administratives et les ortho-photos), un état de l'art et un résumé des contrôles effectués par les agences cartographiques internationales ont été dressés dans la même optique que pour les trois premières couches vectorielles

Critère/Sous-critère	Méthode ou outil conseillé		
	Bâtiments	Axes de voirie	Points d'adresses
Cohérence logique	Logiciel SIG/ETL	Logiciel SIG/ETL	Logiciel SIG/ETL
Excédent	Ortho-photos	Ortho-photos	ICAR
Omission	Ortho-photos	Ortho-photos	ICAR
Positionnement	Terrain	Terrain	Terrain
Classement des objets	Ortho-photos + Schéma conceptuel	Ortho-photos + Schéma conceptuel	Ortho-photos + Schéma conceptuel
Classement/évaluation des attributs	Schéma conceptuel	Schéma conceptuel	/
Qualité temporelle	Logiciel SIG/ETL	Logiciel SIG/ETL	Logiciel SIG/ETL
Généalogie	Logiciel SIG/ETL	Logiciel SIG/ETL	Logiciel SIG/ETL

**Tableau 2.** Méthodes et outils de validation selon les usages internationaux. ICAR = Inventaire Centralisé des Adresses et des Rues en Wallonie ; SIG = Système d'information géographique ; ETL = *Extract / Transform / Load*.

(bâtiments, axes de voiries et points d'adresses). Sans entrer dans une étude approfondie des différentes couches, des prescriptions et des recommandations ont été fournies pour permettre une étude *a posteriori*.

#### IV. AMÉLIORATION CONTINUE

L'amélioration continue des processus de création et de mise à jour des données géographiques est une démarche complexe et qui doit inévitablement s'inscrire dans un projet stratégique à long terme. La Direction de la Géométrie est en charge de ce processus. De ce fait, elle coordonne et réalise des campagnes de contrôle des données géographiques sur l'entièreté du territoire wallon. Ce contrôle est réalisé au gré des mises à jour afin de garantir le niveau de qualité dans le temps.

On l'a dit, ce processus est long et nécessite d'importantes ressources en personnel, en temps, en argent, etc. De façon à mieux faire correspondre l'offre aux besoins évolutifs des utilisateurs de données, il est primordial de vérifier que la proposition de services évolue elle aussi. En réalité, c'est ce dernier point qui est le plus contraignant : l'évaluation du progrès à mettre en œuvre.

À l'inverse, les outils évoluent eux aussi : les techniques de mesures s'améliorent, le nombre de fonctionnalités des différents programmes augmente, l'expérience du personnel s'accroît, etc. Cependant, les fournisseurs de solutions informatiques ne sont pas toujours au plus proche de leurs utilisateurs et

il est fort probable, par conséquent, que les améliorations des premiers n'aillent pas à la même vitesse que les besoins des seconds, et inversement.

Ce processus itératif de conception est traduit en quatre actions qui se succèdent logiquement : la planification, le développement, le contrôle et l'action (PDCA, ou « *plan, do, check, act* » en anglais) (voir Figure 2). Naturellement, la phase d'entrée dans ce cycle est la phase de planification. Cette phase est elle-même divisée en trois sous-parties :

- identification du problème à résoudre : produire de l'information géographique et en garantir la qualité ;
- recherche des causes : quels sont les besoins des utilisateurs de cette information géographique en matière de précision, de cohérence, de complétude, etc. ;
- recherche de solutions et établissement d'un planning : il s'agit en réalité de rédiger le cahier des charges.

Sur un plan stratégique, la phase de planification correspond à l'évaluation des besoins. Ensuite, on agit : les équipes de cartographes améliorent la complétude des données spatiales et des différentes couches d'information. Ils effectuent les mises à jour au gré des missions des équipes de terrain et des partenaires du service public. C'est la phase de développement.

Dans un troisième temps, celui du contrôle, il peut être nécessaire de revoir les besoins mis en évidence lors de la première phase. Cette étude de

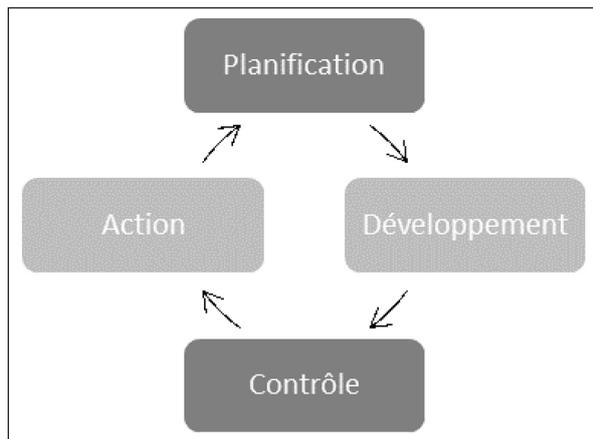


Figure 2. Cycle PDCA

contrôle de la qualité peut se faire, tantôt de façon interne par les équipes au fur et à mesure de la production de nouvelles informations, tantôt par un organisme indépendant externe. Le projet QualiGeo s'inscrit dans cette seconde hypothèse : on vérifie, dans le cadre du projet, le contrôle de la qualité.

La quatrième phase, celle de l'action, est la suite du projet. Il s'agit alors d'appliquer les propositions du projet et de reprendre le cycle en tenant compte des nouvelles connaissances acquises lors des étapes précédentes. Une nouvelle étape de planification démarre alors.

## CONCLUSION

La première étape de toute étude de qualité devrait commencer par la confrontation du modèle conceptuel de données avec une réévaluation des besoins des utilisateurs. La traduction de ces besoins n'est pas toujours aisée ou directe mais elle est essentielle. Outre ces exigences, les choix qui ont guidé la mise en place de ce modèle sont bien souvent tributaires de la stratégie globale mise en place par le gestionnaire de la donnée. Il est donc aussi essentiel que ce gestionnaire fasse une publicité claire et complète du dit modèle et, parallèlement, de la volonté sous-jacente transcrite dans ces choix.

La mise à jour et le suivi de la qualité des données, mais aussi de leur modèle conceptuel, représentent un travail complexe. Un retour est nécessaire sur ce modèle à échéance plus ou moins régulière pour statuer sur l'évolution des besoins des utilisateurs.

L'étude sur le contrôle de la qualité présentée ici répond à la demande formulée par la directive

INSPIRE en matière de documentation de la qualité mais aussi pour la révision du modèle de données initial. Elle fournit un cadre méthodologique destiné tant au gestionnaire de la donnée qu'aux prestataires qui pourraient être chargés de sa maintenance, de sa mise à jour, de son enrichissement, etc. Cette étude est une base pour la rédaction d'un cahier des charges spécial mais n'en constitue pas un en soi. Pour rappel, les spécifications reprises dans le document sont des spécifications de résultats et non pas de moyens. Il est important, au vu des dernières avancées tant stratégiques que techniques, de ne pas limiter les possibilités et l'innovation comme la tendance actuelle semble le conduire.

## BIBLIOGRAPHIE

- Arrêté du 16 septembre 2003 portant sur les classes de précision applicables aux catégories de travaux topographiques réalisés par l'Etat, les collectivités locales et leurs établissements publics ou exécutés pour leur compte. (2003). Journal Officiel 'Lois et Décrets'. <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000000794936>
- Bel Hadj Ali, A., & Vauglin, F. (2000, July). Assessing positional and shape accuracy of polygons in vector GIS. Proceedings of the 4th International Symposium on Spatial Accuracy. 4th International Symposium on Spatial Accuracy, Amsterdam, Netherlands.
- David, B., & Fasquel, P. (1997). Qualité d'une base de données géographique: Concepts et terminologie (No. 53). Institut géographique national. <https://hal-lara.archives-ouvertes.fr/hal-02372984>
- Décret relatif à l'infrastructure d'information géographique wallonne. (2010). Parlement Wallon (Wallex). <https://wallex.walloonie.be/eli/loi-decret/2010/12/22/2011200519/2018/07/17>
- Delavar, M. R., & Devillers, R. (2010). Spatial Data Quality: From Process to Decisions: Guest Editorial. Transactions in GIS, 14(4), 379–386. <https://doi.org/10.1111/j.1467-9671.2010.01224.x>
- Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). (2007). Official Journal of the European Union (EurLEX). <https://eur-lex.europa.eu/eli/dir/2007/2/2019-06-26>
- Guptill, S. C., & Morrison, J. L. (1995). Elements of Spatial Data Quality (Vol. 1–1). Elsevier. <https://doi.org/10.1016/C2009-0-14900-0>
- Hangouët, J.-F. (1995, February 27). Computation of the Hausdorff Distance between plane vector polyline. 12th AutoCarto Conference, Charlotte, USA.
- International Organization for Standardization. (2013). Geographic information—Data quality (ISO Stan-

dard ISO 19157:2013). <https://www.iso.org/obp/ui/#iso:std:iso:19157:ed-1:v1:en>

Plan opérationnel géomatique pour la Wallonie 2017-2019. (2017). Service Public de Wallonie - Département de la Géomatique. [https://geoportail.wallonie.be/files/documents/POGW\\_2017\\_2019.pdf](https://geoportail.wallonie.be/files/documents/POGW_2017_2019.pdf)

Vauglin, F. (1999). A practical study on precision and resolution in vector geographical databases. Proceedings of International Symposium on Spatial Data Quality, 10.

### *Coordonnées des auteurs :*

Gilles-Antoine NYS  
Unité de Géomatique  
Département de Géographie  
UR SPHERES, Université de Liège  
Auteur correspondant : ganys@uliege.be

Charline DUBOIS  
Unité de Géomatique  
Département de Géographie  
UR SPHERES, Université de Liège  
charline.dubois@uliege.be

Céline GOFFIN  
Unité de Géomatique  
Département de Géographie  
UR SPHERES, Université de Liège  
celinegoffin@live.fr

Pierre HALLOT  
DIVA, UR AAP, Université de Liège  
p.hallot@uliege.be

Jean-Paul KASPRZYK  
Unité de Géomatique et SEGEFA  
Département de Géographie  
UR SPHERES, Université de Liège  
JP.Kasprzyk@uliege.be

Mathieu TREFFER  
Unité de Géomatique  
Département de Géographie  
UR SPHERES, Université de Liège

Roland BILLEN  
Unité de Géomatique  
Département de Géographie  
UR SPHERES, Université de Liège  
rbillen@uliege.be