



HAL
open science

FIABILISATION DES RECUEILS DE DONNÉES D'ENQUÊTES OD ROUTIÈRES

Fabrice Hasiak, Bertrand Zogall

► **To cite this version:**

Fabrice Hasiak, Bertrand Zogall. FIABILISATION DES RECUEILS DE DONNÉES D'ENQUÊTES OD ROUTIÈRES. 7ème Colloque francophone sur les sondages, Nov 2012, Rennes, France. <hal-01739720>

HAL Id: hal-01739720

<https://hal.science/hal-01739720v1>

Submitted on 21 Mar 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

FIABILISATION DES RECUEILS DE DONNÉES D'ENQUÊTES OD ROUTIÈRES

Fabrice HASIAK¹ & Bertrand ZOGALL¹

¹ *Cete Nord Picardie (CETE NP), 2 rue de Bruxelles BP 275 – 59019 LILLE Cedex*

fabrice.hasiak@developpement-durable.gouv.fr

bertrand.zogall@developpement-durable.gouv.fr

Résumé

La qualité des enquêtes sur les déplacements des individus, de manière générale, dépend à la fois du mode de passation des questionnaires (enquêtes en face à face, auto-administrées) [Cete NP, 2004] et de la manière dont le suivi aura pu être fait pendant toute la phase de l'enquête, depuis sa préparation jusqu'à la fourniture des données recueillies. C'est particulièrement vrai sur certains types d'enquêtes comme les enquêtes Origine Destination (enquêtes OD) routières réalisées par interviews des conducteurs de véhicules en face à face où :

- les personnes chargées de réaliser ces enquêtes sont surtout des personnes faisant ce travail d'enquêteur de façon très occasionnelle, comme « boulot » d'appoint (étudiants, retraités, personnes à la recherche d'un emploi), et donc peu familiarisées aux objectifs de ce type d'enquête et aux vocabulaires employés dans les questionnaires;
- les enquêteurs sont en permanence sous pression, car ce sont des enquêtes très courtes (moins de 3 minutes), nécessitant de leur part une forte réactivité pour faire face à certaines situations difficiles (mécontentement des usagers enquêtés, incompréhension des questions, ...).

Cet article se propose de faire le point sur les méthodes développées au Cete NP pour améliorer la qualité du recueil des déplacements dans les enquêtes OD: formulation des questions dans les questionnaires, suivi et contrôle terrain des interviews [Certu, 2008] et [Setra, 2010] et apurement des données. Sur ce dernier point, le Cete NP a expérimenté et mis au point plusieurs méthodes pour fiabiliser les informations sur les OD. Il s'agit de vérifier que l'OD déclarée est cohérente avec le lieu et le sens d'enquête. Compte tenu des volumes et de la diversité des OD enquêtées, les méthodes développées s'appuient en partie sur des processus de contrôles automatiques: soit à partir d'outils de type logiciel d'affectation de trafic (mais ces outils coûteux n'existent pas toujours sur le territoire enquêté), soit à partir de concepts purement mathématiques (facilement reproductibles), s'appuyant notamment sur le calcul et la comparaison des distances OD directes et via le poste d'enquête, et l'analyse de l'angle formé par l'origine, le poste d'enquête et la destination. Les tests réalisés permettent en moyenne de contrôler automatiquement la moitié des OD enquêtées sur un poste d'enquête. L'autre moitié des OD est contrôlée « à la main » en opérant un regroupement des origines et des destinations dites « semblables » selon le principe suivant: plus un point donné est éloigné du poste d'enquête plus il peut être regroupé avec d'autres points.

Les enquêtes OD routières : une méthodologie bien rodée

Les enquêtes OD routières, qui existent depuis plus de 30 ans, sont des enquêtes réalisées en face à face auprès d'un échantillon de conducteurs de véhicule (voitures et poids lourds, français et étrangers). L'objectif de ces enquêtes est de mieux connaître les flux de véhicules circulant sur un ou plusieurs axes. Pour ce faire, les véhicules peuvent être arrêtés de deux façons :

- Cas 1 : soit par les forces de l'ordre. Dans ce cas, les véhicules sont orientés vers des aires d'enquête hors des voies de circulation.
- Cas 2 : soit au moyen de feux de chantier. Dans ce deuxième cas, les véhicules sont directement enquêtés sur une voie de circulation.

Pour ce type d'enquête, les questionnaires sont très courts et la durée des interviews n'excède pas 3 minutes dans le premier cas et 45 secondes dans le second (ce qui correspond à une durée de cycle moyen de feu en agglomération).



Illustration 01: Arrêt des véhicules hors des voies de (source : CETE Nord Picardie)



Illustration 02: Arrêt des véhicules sur voie de circulation avec feux de chantier (source : CETE Nord Picardie)

La qualité des enquêtes OD: une préoccupation permanente, depuis la préparation amont des enquêtes jusqu'à la fourniture finale des données

Aujourd'hui, la qualité du recueil des enquêtes OD (comme bon nombre d'enquêtes de mobilité, en particulier les enquêtes ménages déplacements en particulier [CERTU, 2008]) est assurée à plusieurs niveaux: en amont avant le lancement de l'enquête, pendant et après l'enquête.

En amont des enquêtes sur le terrain, la phase de formation des enquêteurs et du personnel encadrant est primordiale. En général, les enquêteurs suivent 2 à 3 heures de formation avec mises en situation (le plus souvent en bureau, dans le meilleur des cas sur le terrain) avant de réaliser réellement les enquêtes OD. Cette formation est essentielle pour permettre une bonne appropriation du questionnaire par les futurs enquêteurs. Elle permet aussi de repérer les personnes qui ne seront pas capables de faire les interviews.

Sur le terrain ensuite, l'encadrement des enquêteurs par un chef d'équipe doit également permettre d'assurer un bon recueil de données. Le rôle du chef d'équipe, au-delà de veiller au bon fonctionnement du dispositif d'enquête (dont la sécurité des enquêteurs sur la route, les contacts avec les forces de l'ordre pour la gestion des échantillons à enquêter), est de suivre la qualité du recueil par les enquêteurs (bonne formulation des questions, relevé correct des informations données par les enquêtés, respect du protocole de passation des enquêtes). Régulièrement, un contrôle externe du prestataire terrain en charge des enquêtes est réalisé en plus, soit directement par le maître d'ouvrage de l'opération, soit par un prestataire extérieur (comme le Centre d'Etudes Techniques de l'Équipement lors d'enquêtes sous maîtrise d'ouvrage de l'Etat). Le service en charge de ce contrôle vérifiera le bon déroulement des enquêtes: contrôle du personnel encadrant et du travail des enquêteurs. Ce processus de contrôle est indispensable pour éviter toute dérive et tout aménagement du protocole de passation des enquêtes trop souvent réalisé par les enquêteurs comme semble le montrer une récente recherche du sociologue Rémy Caveng [Caveng, 2012].

Enfin, une fois l'enquête terminée, lors de la constitution de la base de données d'enquête, un dernier contrôle qualité peut être effectué. Ce contrôle (le plus souvent assez simple) consiste à interdire certaines saisies des données (modalités de réponse non prévue) et/ou à vérifier la cohérence entre certaines réponses d'un même questionnaire, ce type de procédure pouvant être directement intégré dans les questionnaires si la passation des enquêtes lors de la phase terrain se fait à l'aide de terminaux informatiques portables.

Malgré tous les contrôles, des erreurs dans les questionnaires peuvent persister

Malgré toutes les précautions prises pour assurer une bonne qualité du recueil de donnée (formation des enquêteurs, encadrement terrain, contrôle externe...), il n'est pas rare de constater dans les fichiers d'enquêtes transmis par les prestataires d'enquête des erreurs dans les données sur les déplacements (origine - destination) recueillis :

- OD impossible par rapport à la localisation du poste d'enquête ;
- OD recueillie mal orientée par rapport au sens de l'enquête.

Ces erreurs sur OD peuvent provenir :

- de l'enquêteur : par une mauvaise formulation des questions (et donc mauvaise compréhension de l'enquêté entraînant une mauvaise réponse) ou par une mauvaise retranscription par l'enquêteur des informations données par l'enquêté.
- de l'enquêté : par mauvaise compréhension des questions posées par l'enquêteur, par une réponse incomplète de l'enquêté (exemple « *je vais à Aix* » mais quel Aix ? Aix-en-

Provence/Aix-Noulette/Aix d'Angillon...).

- des opérateurs de codification et de saisie des enquêtes : par une retranscription erronée des données recueillies sur les questionnaires.

En général, on constate aujourd'hui, sur un cordon d'enquêtes OD, en moyenne, 5% d'OD incorrectes dans les bases de données (ce taux pouvant être plus important lorsque la formation des enquêteurs est trop courte ou lorsque le suivi qualitatif des enquêteurs sur le terrain est insuffisant). Ces erreurs ne sont pas toujours « corrigées » (cas des OD inversées par rapport au sens de l'enquête) ou supprimées (cas des OD impossibles ou incohérentes par rapport à la localisation du poste d'enquête) car elles ne sont pas facilement détectables au moment de la réception des fichiers d'enquête. Très souvent donc, elles sont détectées au moment de l'exploitation des données et en particulier lors de la production de matrice de flux OD.

Comment mieux fiabiliser encore les données des enquêtes OD ?

Afin d'améliorer encore les recueils d'enquêtes OD routières, nous avons exploré plusieurs pistes. Ces pistes concernent :

- une nouvelle mise en forme des questionnaires et une nouvelle formulation des questions ;
- une systématisation du contrôle extérieur de la maîtrise d'œuvre ;
- un contrôle approfondi, au moment de l'apurement du fichier d'enquête, des OD.

Une contrainte forte des enquêtes OD, enquêtes administrées en face à face, est la durée d'interview qui est très réduite (3 minutes dans le meilleur des cas). Il faut donc que l'enquêteur se fasse comprendre très vite par les enquêtés. La formulation des questions aux enquêtés et l'ordre des questions doivent donc être très proches d'une conversation. Or les questionnaires habituellement utilisés pour les enquêtes OD restent encore trop « techno » et utilisent des concepts qui ne peuvent être compris par tous (exemple : « *quelle est l'origine de votre déplacement ?* », « *quel est votre motif origine ?* »...). Certes, un enquêteur bien formé pourra facilement se faire comprendre et reformuler éventuellement une question mal comprise. Toutefois les enquêteurs sont rarement « performants » dès le début des enquêtes (il y a toujours un temps d'adaptation) et rares sont les questionnaires sans erreur. Un travail avec un sociologue non spécialiste des enquêtes OD (Joël Meissonnier, chargé de recherche au CETE Nord Picardie) a permis de restructurer et reformuler les questions d'enquête OD. Le résultat de ce travail est présenté ci-dessous.

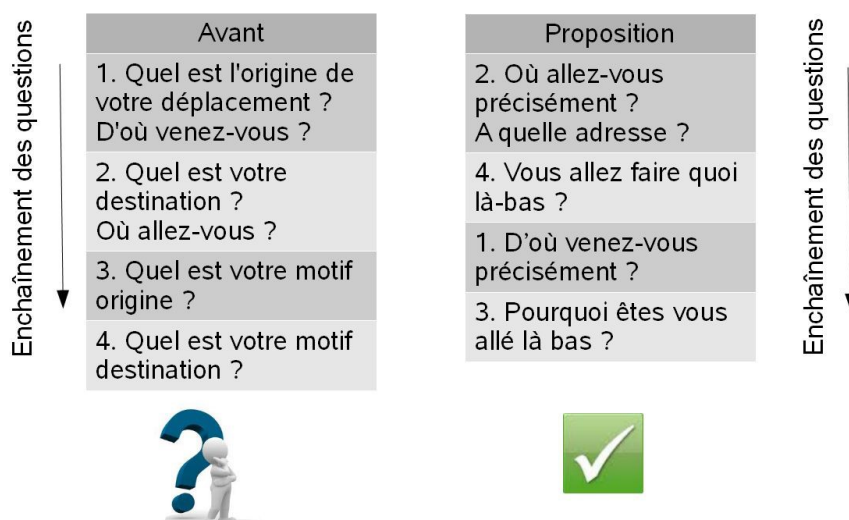


Illustration 03: Proposition d'enchaînement des questions d'enquête OD (source : CETE Nord Picardie)

Le contrôle externe du prestataire d'enquête est aussi un élément important permettant d'assurer une bonne qualité du recueil de données. La plupart des bureaux d'études spécialisés dans les enquêtes mettent en place des procédures internes de contrôle de leurs prestations. Toutefois, un œil extérieur est toujours préférable pour contrôler le bon déroulement des enquêtes tant en amont (veiller à ce que les formations se déroulent dans de bonnes conditions : dans une salle en petits groupes et de durée suffisante (au moins 2 à 3 heures)) que pendant la phase terrain (contrôle du travail du personnel encadrant et des enquêteurs).

Le dernier point concerne l'apurement des fichiers d'enquête. Au-delà des tests « classiques » de contrôle des différentes données présentes dans un fichier d'enquête OD (contrôles des modalités de réponse à une question donnée) généralement faits, rares sont les bureaux d'études qui réalisent en plus un contrôle des OD recueillies et saisies par rapport à la localisation des lieux d'enquête. Il est vrai qu'un tel contrôle n'est pas facile à mettre en place de façon plus ou moins automatique. Ce type de contrôle sur les OD est en général fait à la main relation par relation (les OD commune à commune sont contrôlées par rapport au lieu d'enquête et par rapport au sens de l'enquête). Compte tenu de la très grande diversité des OD recueillies lors d'une enquête de type enquête cordon avec de nombreux axes routiers enquêtés et compte tenu des volumes d'enquête en général réalisés (1000 à 2000 enquêtes par axe), ce type de contrôle « à la main » peut vite se révéler fastidieux. Confronté régulièrement à ce type de travail, nous avons cherché à automatiser le contrôle des OD. La première piste testée a été d'utiliser des logiciels d'affectation de trafic type Emme4 (<http://www.inrosoft.com>) en affectant les matrices de flux OD enquêtés (sur un poste d'enquête) sur un réseau modélisé. En sortie, le logiciel d'affectation permet de visualiser les itinéraires « vraisemblables » des usagers (itinéraires modélisés). En confrontant ces itinéraires avec la localisation du poste d'enquête on peut donc aisément repérer les OD correctes et les OD incorrectes (celles ne passant pas par le poste d'enquête).

Affectation des enquêtes OD Autoroute A1 sur le réseau modélisé

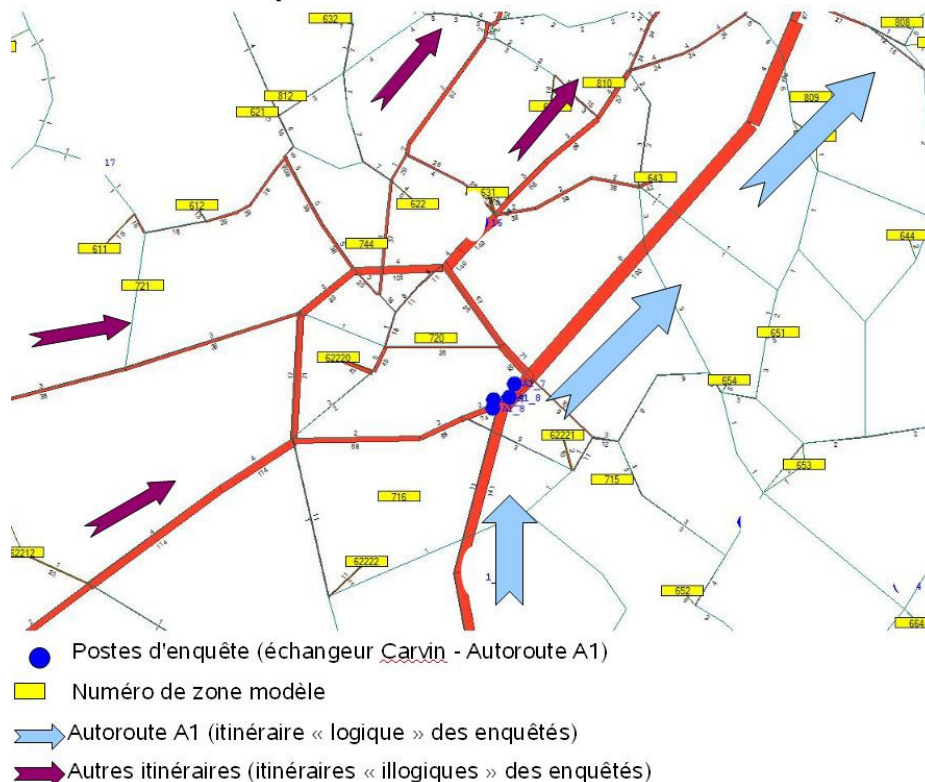


Illustration 04: Itinéraires reconstitués par modèle des flux OD enquêtés sur l'autoroute A1 (source : CETE Nord Picardie)

Cependant, l'utilisation de tels modèles n'est pas aisée et certains problèmes intrinsèques à la modélisation peuvent rendre caduque ce contrôle automatique des OD. En effet, les modèles d'affectation de trafic sont en général calés sur l'heure de pointe alors que les enquêtes OD sont réalisées sur toute une journée. On peut donc être confronté à l'invalidation (automatique) d'une OD liée simplement à un problème de modélisation. Il faut donc être extrêmement prudent quant à l'utilisation de tels outils pour valider ou non les enquêtes OD.

Face à ce constat, nous avons développé une méthode basée à la fois sur un contrôle automatisé des OD et sur un contrôle manuel « simplifié » pour fiabiliser les bases de données d'enquête OD. Cette méthode est facilement dupliquable puisqu'elle ne nécessite pas l'utilisation d'outils professionnels très spécialisés comme les logiciels d'affectation de trafic.

Vers une nouvelle méthode, développée au CETE Nord Picardie, pour le contrôle des OD dans les fichiers d'enquête

L'automatisation (ou semi-automatisation) du contrôle des OD est effectuée à partir du fichier obtenu après la saisie informatique des questionnaires issus de la phase d'enquête terrain. Pour chaque enregistrement (correspondant à un usager enquêté), ce fichier comprend notamment l'origine et la destination du trajet, avec une précision variable (rue, commune, département ou pays). Le processus de contrôle automatique des OD comporte plusieurs phases.

Phase 0 : Préparation des données d'enquête - géolocalisation des origines et des destinations

La première étape consiste à géolocaliser (sous SIG) les origines et destinations du fichier d'enquête comportant soit un code INSEE (pour les communes françaises) ou une autre codification (par exemple le code province ou région pour la Belgique ou le code pays pour les pays Étrangers). Cette géolocalisation permet ainsi d'obtenir pour chaque relation, les coordonnées (x,y) des origines et des destinations des déplacements enquêtés. Le poste d'enquête (P) est également géolocalisé.

Tous les déplacements pour lesquels il y a une imprécision sur l'une des extrémités (O ou D), par conséquent non géolocalisable, sont exclus préalablement du fichier, car de toute façon inexploitable.

Étape 1 : Contrôle automatique des OD à l'aide de critères géométriques

Les coordonnées (x,y) de chaque O et D et celles du poste d'enquête et l'application de théorèmes mathématiques (notamment le théorème d'Al-Kashi) permettent de calculer plusieurs données et indicateurs mathématiques qui serviront à la validation automatique ou non des flux enquêtés:

- les distances à vol d'oiseau [Origine-Destination], [Origine-Poste] et [Poste-Destination] ;
- le rapport [distance Origine-Destination / (distance Origine-Poste + distance Poste-Destination)] noté Div_dist;
- l'angle formé par l'origine, le poste et la destination noté Angle (OpD);
- l'angle [direction principale-poste-destination] noté Angle (DPpD), la direction principale correspondant à « l'orientation » de l'axe enquêté (cette direction dépend du sens enquêté).

Ces différents éléments permettent de déterminer si l'OD recensée est compatible avec le passage par le poste d'enquête.

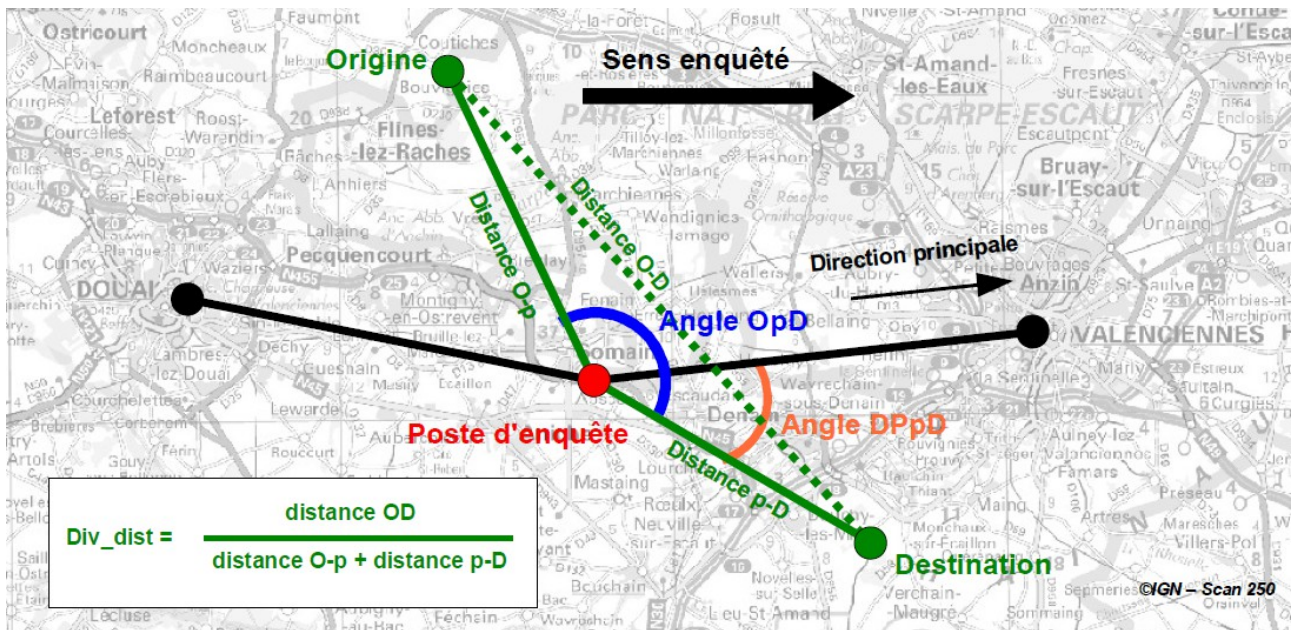


Illustration 05: exemple d'application des critères géométriques aux contrôles des OD (source : CETE Nord Picardie)

Pour les trajets dont l'origine et la destination sont géolocalisées, il s'agit alors de définir des critères qui vont permettre d'attribuer automatiquement un « code contrôle » et ainsi valider ou non l'OD enquêtée:

- « 0 » pour une OD à rejeter (=OD non correcte par rapport à la localisation du poste d'enquête);
- « 1 » pour une OD à conserver (=OD validée);
- « 2 » pour une OD à inverser (=OD à inverser par rapport au sens de l'enquête);
- « manuel » pour une OD à traiter manuellement (ce sont les OD pour lesquelles les critères géométriques ne permettent pas de déterminer automatiquement si elles sont à rejeter, conserver ou inverser).

Ainsi, une OD sera considérée comme valide si le rapport des distances Div_dist est « proche de 1 » et si l'angle (OpD) est « plutôt plat ». Le dernier critère Angle (DPpD) permet de valider ou corriger le sens de l'OD par rapport au sens de l'enquête. Le sens de l'OD sera considéré comme correct si cet angle (DPpD) est plutôt « proche de zéro ».

Pour définir ces critères (c'est-à-dire des valeurs limites des éléments géométriques définis plus haut), nous avons d'abord traité manuellement toutes les OD d'un poste avant de confronter les résultats avec les éléments géométriques. Par exemple, on remarque que pour toutes les OD codées « 0 », l'angle (OpD) est inférieur à 90 degrés. Pour éviter de définir des critères trop hasardeux, les origines et destinations trop proches du poste d'enquête (à moins de 5km) et les longs trajets (plus de 150 km) ont été exclus. Ces OD sortent du processus de contrôle automatique et doivent faire l'objet d'un contrôle manuel. Les différents critères choisis sont présentés ci-après.

Critères utilisés pour le poste de l'A21

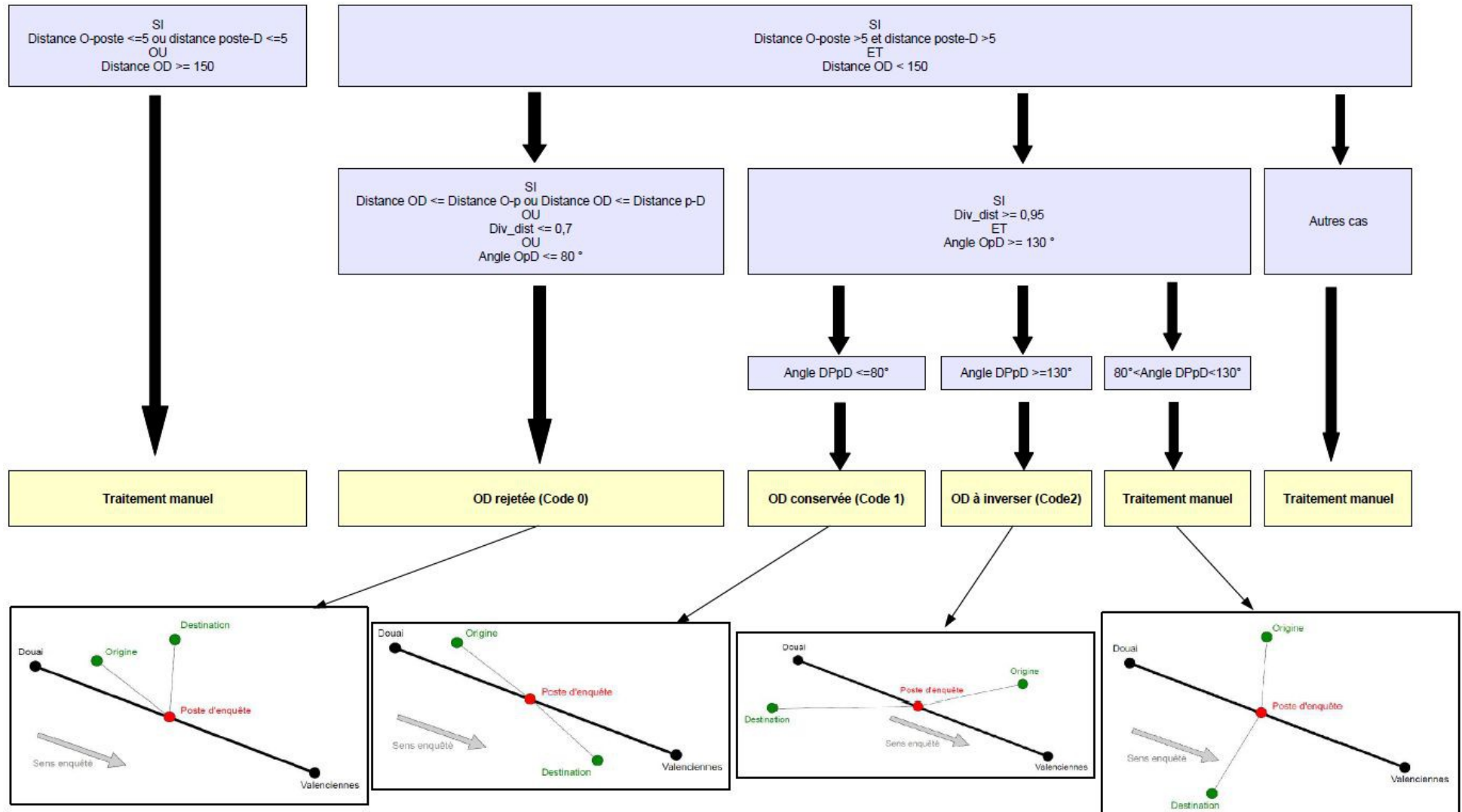


Illustration 06: Critères géométriques utilisés pour le contrôle du poste d'enquête sur l'autoroute A21 entre Douai et Valenciennes (source : CETE Nord Picardie)

Étape 2 : Contrôle manuel par regroupement des OD

A l'issue de l'étape 1, toutes les OD n'ont pas pu être validées (ou invalidées) de manière automatique. Cette étape 2 consiste à effectuer des regroupements géographiques des OD restant à traiter, afin de les contrôler manuellement « par paquet » et non commune à commune. A l'aide du logiciel Mapinfo, on attribue donc à chaque origine et chaque destination une zone, dont la finesse dépend de sa distance par rapport au poste d'enquête. Par exemple, pour une extrémité située à moins de 5 km du poste, la zone attribuée sera la commune (niveau le plus fin). Plus on s'éloigne du poste d'enquête, plus on procédera à des regroupements des origines et des destinations. Ce zonage peut utiliser les limites administratives (canton, arrondissement, département, région...) ou être spécifique. Il doit être adapté en fonction du poste d'enquête. Il s'agit alors de contrôler les OD regroupées selon le zonage défini. Ainsi, si plusieurs trajets ont leur origine dans la zone « A » et leur destination dans la zone « B », il suffit de contrôler l'OD « A-B » et d'attribuer le code contrôle adéquat à tous ces trajets. Dans le cas des enquêtes réalisées dans les deux sens de circulation, pour réduire encore le nombre d'OD à contrôler, on regroupe les OD « symétriques ». En effet, on sait que le trajet « A-B » dans le sens 1 aura le même code contrôle que l'OD « B-A » dans le sens 2. Il suffit donc d'inverser tous les trajets du sens 2 (par exemple) et de les contrôler comme s'ils étaient dans le sens 1.

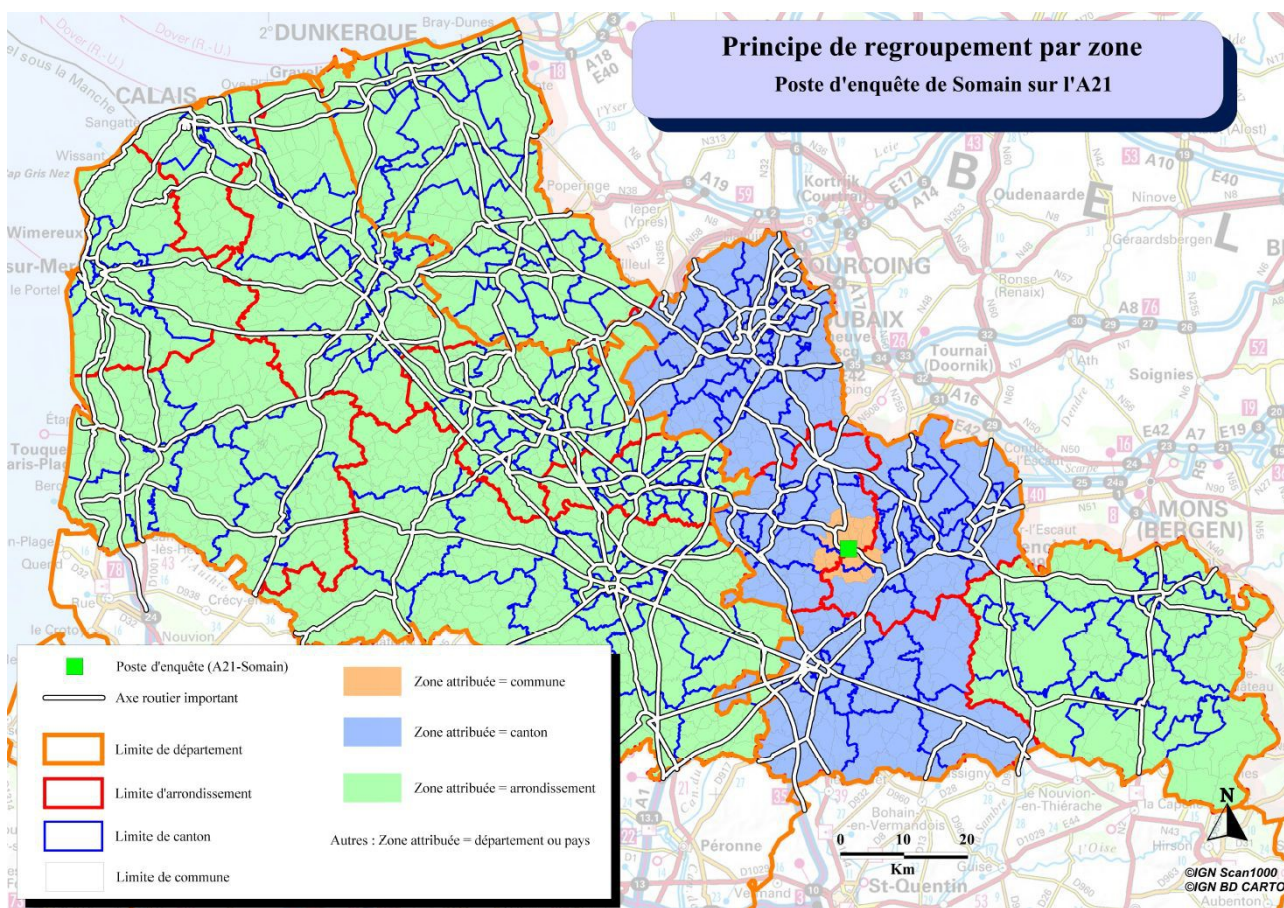


Illustration 07: Principe de regroupement par zone – exemple sur l'A21 (source : CETE Nord Picardie)

Résultats

Dans le cas de l'enquête sur l'autoroute A21 à Somain, le contrôle concernait 3333 trajets enquêtés, correspondant à 2014 OD différentes.

La présente méthode a permis d'obtenir les résultats suivants :

- 60 OD avec une extrémité inconnue dans la région Nord-Pas-de-Calais ou en Belgique ont été automatiquement codées 0 (OD rejetées) (Phase 0);
- 1644 trajets ont été traités automatiquement, à l'aide de critères géométriques (étape 1) ;
- le regroupement des OD fines a permis de ne traiter que 462 « OD entre zones », correspondant à 1629 déplacements (étape 2).

CODE contrôle	Phase 0		Etape 1 Contrôle automatique		Etape 2 Contrôle manuel		Synthèse	
	Nombre	%	Nombre	%	Nombre	%	Nombre	%
0 (OD rejetée)	60	1,8%	74	2,2%	138	4,1%	272	8,2%
1 (OD valide)			1492	44,8%	1431	42,9%	2923	87,7%
2 (OD à inverser)			78	2,3%	55	1,7%	133	4,0%
9 (OD indécode)			0	0,0%	5	0,2%	5	0,2%
Total	60	1,8%	1644	49,3%	1629	48,9%	3 333	100,0%

[le code 9 correspond à des trajets pour lesquels il est difficile de se prononcer]

L'application de ce contrôle « semi-automatisé » des enquêtes OD permet un gain de temps non négligeable. Celui-ci est estimé à 50% par rapport à un contrôle totalement manuel.

Conclusion

L'ensemble des méthodes évoquées ci-dessus permettent d'améliorer la fiabilité des données recueillies lors des enquêtes Origine-Destination routières tout en limitant les opérations de contrôle. Ces méthodes interviennent à toutes les phases de réalisation d'une enquête OD: avant, pendant et après l'enquête de terrain.

Avant l'enquête de terrain, il s'agit de former des enquêteurs performants et de leur fournir les moyens d'assurer leur mission efficacement. Cela passe par la conception d'un questionnaire d'enquête adapté et par une bonne appropriation de ce questionnaire par les enquêteurs.

Pendant l'enquête, il s'agit d'assurer un recueil de données le plus fiable possible. Un encadrement efficace des enquêteurs, notamment en début d'enquête, et un contrôle régulier des données recueillies sont nécessaires à l'atteinte de cet objectif.

Enfin après l'enquête terrain, il convient de vérifier la cohérence des données recueillies. Les méthodes présentées ici permettent de s'assurer de la fiabilité des données en limitant le temps consacré à l'apurement.

Bibliographie

- Caveng R. (2012). La production des enquêtes quantitatives. Revue d'anthropologie des connaissances. 2012. 25 p.
- Cete NP (2004). Étude méthodologique sur la connaissance des déplacements des périurbains. Tome 2. Méthodologie d'enquête sur les déplacements des périurbains. CERTU et DDE59. 2004. 68 p.
- Certu (2008). L'enquête ménages déplacements « standard Certu ». Guide méthodologique. Collection Références n°72. 2008. 207 p.
- Setra (2010). Guide méthodologique. Enquêtes de circulation Organisation et déroulement. 2010. 104 p.